



Création de Systèmes d'Intelligence dans une Organisation de Recherche et Développement avec la Scientométrie et la Médiométrie

Roberto Penteadó

► **To cite this version:**

Roberto Penteadó. Création de Systèmes d'Intelligence dans une Organisation de Recherche et Développement avec la Scientométrie et la Médiométrie. domain_stic.bibl. Université du Sud Toulon Var, 2006. Français. <tel-00112626>

HAL Id: tel-00112626

<https://tel.archives-ouvertes.fr/tel-00112626>

Submitted on 9 Nov 2006

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

UNIVERSITÉ DU SUD, TOULON-VAR

N° attribué par la bibliothèque

|_|_|_|_|_|_|_|_|_|_|_|_|_|_|_|

THÈSE POUR L'OBTENTION DU DOCTORAT EN SCIENCES DE L'INFORMATION ET DE LA
COMMUNICATION À L'UNIVERSITÉ DU SUD, TOULON-VAR
conforme au nouveau régime défini par l'arrêté du 30 mars 1992

**CRÉATION DE SYSTÈMES D'INTELLIGENCE
DANS UNE ORGANISATION DE RECHERCHE ET
DÉVELOPPEMENT AVEC LA
SCIENTOMÉTRIE ET LA MÉDIAMÉTRIE**

présentée et soutenue publiquement le 10 octobre 2006
par Roberto de Camargo PENTEADO Filho
Entreprise Brésilienne de Recherche Agronomique - Embrapa

Sous la Direction de

M. Luc QUONIAM

Professeur à l'Université du Sud, Toulon-Var

Membres du Jury :

M. Edson Luiz RICCIO

Professeur à l'Université de Sao Paulo

M. Tony TSCHAEGLER

Professeur à l'Université de Nice

M. Henri DOU

Professeur à l'Université de Marseille III

M. Eric BOUTIN (Co-Directeur de Recherche)

Maître de Conférences à l'Université du Sud, Toulon-Var

Résumé

Ce travail est un trait d'union entre les sciences de l'information et de la communication. Une robuste méthodologie et des outils performants d'analyses bibliométriques sont utilisés pour des études scientométriques et médiométriques. Pour cela, nous avons étudié la production scientifique d'une organisation publique de recherche et développement, l'Entreprise Brésilienne de Recherche Agronomique (Embrapa), les compétences de ses chercheurs et enfin nous avons évalué la performance de cette organisation et ses 40 centres de recherche dans les médias.

Les résultats indiquent que la fonction d'analyse d'informations internes et externes aux organisations est aujourd'hui un instrument important et stratégique pour améliorer les décisions et les politiques organisationnelles et contribuer au développement de la société.

L'étude des compétences de R&D démontre le potentiel d'intelligence qui peut être générée avec des informations déjà présentes dans les organisations, mais dispersées en plusieurs départements administratifs. Il est le témoin de la nécessité d'organiser des environnements informatiques spécifiques pour la fonction d'analyse dans les organisations, intégrant bases de données et informations d'origines diverses.

Avec les méthodologies de médiométrie les fonctions d'audit du travail de communication organisationnelle, d'évaluation du résultat de campagnes et de la production de dépêches pour la presse gagnent en efficacité et l'élaboration de suggestions de reportages peut mieux correspondre aux préférences des éditeurs. Ceci ne serait pas possible sans l'acquisition de l'intelligence sur les médias pour déterminer les tendances, les préférences des divers véhicules et le comportement de leurs éditeurs.

La fonction de clipping, d'accompagner ce qui sort dans la presse est présente dans les grandes organisations. L'apport de ce travail consiste à intégrer la fonction analyse d'une base Clipping. A partir de tels corpus et munis des méthodologies et outils décrits plus loin, il est possible de réaliser des études prospectives sur l'évolution historique d'un thème, d'un produit, d'un politicien ou d'une entreprise dans les médias et de faciliter des activités de veille de l'environnement, des concurrents ou de tout un secteur de l'économie.

Abstract

This work is an exercise of union between the sciences of information and of communication. A robust methodology and tools of bibliometric analysis are used for scientometric and mediametric studies. For this, we have studied the scientific production of a public organization of research and development, the Brazilian Agricultural Research Corporation (Embrapa), their researchers competencies and finally we have evaluated the performance of this organization and its 40 research centers in the media.

The results indicate that the function of the analysis of organizational internal and external information is today an important and strategic instrument for the improvement of decisions and to build organizational policies that contribute to the development of the society.

The study of R&D competencies demonstrates the potential of intelligence, which can be generated with information already available in organizations, although dispersed into several administrative departments. It witnesses the need to organize in the organization an informational environment specific for the function of analysis, integrating databases and information from various sources.

The use of mediametry methodologies provides space and efficiency to the monitoring functions of the organizations communication processes, for the evaluation of campaigns and the production of press releases and other communication material. The relationship with the Press also benefits with the elaboration of suggestions of adequate reports to the editor's preferences. This would not be possible without the acquisition of intelligence on the media to determinate its trends through the analysis of its production, the themes preferred by each vehicle and the behavior of their editors.

The function of clipping, following what the Press is publishing is present in organizations of the five continents, the integration of the function of analysis by the creation of the Clipping database is the differential presented by this work. With similar bases, methodologies and tools as those described here, it is possible to carry out prospective studies on the historical evolution of a topic, a product, a politician or a company in the media and to facilitate activities of monitoring the environment, competitors or a whole sector of the economy.

Resumo

Este trabalho é um exercício de união entre os campos da Ciência da Informação e da Ciência da Comunicação. Uma robusta metodologia e ferramentas de análise bibliométrica foram utilizados para estudos de cientometria e de midiametria. Para isso, estudamos a produção científica de uma organização pública de pesquisa e desenvolvimento, a Empresa Brasileira de Pesquisa Agropecuária (Embrapa), as competências de seus pesquisadores e finalmente avaliamos a performance desta organização e de suas 40 unidades descentralizadas na mídia.

Os resultados indicam que a função de análise de informações organizacionais internas e externas é hoje um instrumento importante e estratégico para a melhoria das decisões e para construir políticas organizacionais que contribuam para o desenvolvimento da sociedade.

O estudo das competências de P&D demonstra o potencial da inteligência, que pode ser gerada com informações já disponíveis nas organizações, apesar de estarem dispersas em diversas seções e departamentos administrativos. Ele testemunha a necessidade de organizar na organização um ambiente informacional específico para a função de análise, integrando bases de dados e informações de diversas fontes e origens.

O uso de metodologias de midiametria abre espaço e tornam mais eficientes as funções de monitoramento dos processos de comunicação organizacional, para a avaliação de campanhas e a produção de *press-releases* e outros materiais de comunicação. As relações com a Imprensa também se beneficiam com a elaboração de sugestões de pauta mais adequadas e afinadas com as preferências dos editores. Isto não seria possível sem a aquisição de inteligência sobre a mídia determinando suas tendências por meio da análise de sua produção, dos temas preferidos por cada veículo e do comportamento de seus editores.

A função de Clipping, de acompanhar o que sai na imprensa, está presente em organizações nos cinco continentes. A integração da função de análise pela criação e a alimentação da Base de Dados de Clipping é o diferencial apresentado por este trabalho. Com bases, metodologias e ferramentas similares às descritas aqui é possível levar adiante estudos prospectivos sobre a evolução histórica de um tema, um produto, um político ou uma empresa na mídia e facilitar atividades de monitoramento do meio ambiente, de competidores ou de um setor inteiro da economia.

Remerciements

Cette thèse ne serait jamais possible sans l'appui et l'orientation dans toutes les heures du professeur Luc Quoniam. Ont contribué aussi les professeurs Henri Dou, Leandro Faria, Eric Boutin et José Angelo Gregolin. À eux, pour me tenir dans leurs cercles intellectuels et conviviaux, parmi lesquels j'ai beaucoup profité et grandi comme être humain, mon plus sincère remerciement.

Aux distingués membres du Jury pour accepter de m'écouter et pour me guider encore une fois.

Au directeur-président de l'Embrapa, Silvio Crestana et au chef de l'ACS, Edilson Fragalle, pour l'appui nécessaire pour dépasser les derniers obstacles.

À l'Université du Sud, Toulon-Var et à son École Doctorale pour m'accueillir pendant ces quatre ans.

À l'Embrapa, au Cendotec, au NIT-UFSCAR, au CRRM, au Laboratoire I3M (ancien Lepont) et à tous leurs collaborateurs pour leur camaraderie et appui.

À mes conseillers académiques, Mariza Barbosa et Flavio Avila et à Veslei Caetano, Leila Lenk, Maria Helena Kurihara, Rosângela Galon, Simara Carvalho et Jorge Reti pour leur support.

Au personnel de l'ACS et en spécial à Ubiraci Gomes, pour rendre possible la matérialisation des idées qui sont devenues le Guide de Sources et le Clipping Électronique.

À ma famille qui n'a pas reçu l'attention qu'elle méritait, pour la patience et la compréhension.

À la France, aux Françaises et Français, pour être la France, les Françaises et les Français.
Mon éternel remerciement.

SOMMAIRE

Introduction	27
Le Plan de Thèse	28
Définitions et synthèses	31
Société de l'information et de la communication	31
Communication et nouvelles technologies	32
Le modèle social	32
Le modèle psychologique.....	33
Le modèle mathématique	34
Le modèle de main double	35
Les fonctions de la communication	36
Les NTC	36
Différences entre masse et public	38
Relations publiques et marketing.....	39
Données, information et systèmes	40
Connaissance et information	41
Le défi de connaître l'environnement.....	42
Intelligence et Veille	45
Dimension économique de l'intelligence.....	48
Dimension sociale de l'intelligence.....	53
Intelligence et communication	57
La nécessité d'information des organisations	60
Morphologie de l'information.....	62
La relation de l'information avec le temps	63
Le cycle de l'intelligence	65
Collection et stockage : Les entrepôts de données.....	66
Préparation de données et de textes : Condition de qualité	68
Un différentiel dans la conversion des données.....	70
Objectif et concept basique.....	71
Caractéristiques.....	71
La construction des dubcodes	72
Analyse et synthèse de l'information	73

L'analyse et l'exploration de données	75
Le cycle d'analyse dans l'exploration de données.....	75
Exemples d'applications intelligentes de l'exploration de données	77
Des couches et de la bière	77
Des bonnes relations avec le client.....	78
Crédit et insolvabilité	78
Générer plus de négoce.....	78
Veiller le marché	79
Des contrats de maintenance.....	79
Facteurs critiques de succès (FCS).....	79
L'analyse et l'exploration de textes	80
La méthode d'analyse bibliométrique.....	83
Les caractéristiques essentielles des informations traitées	84
Les nouvelles frontières.....	87
La qualité en bases de données	88
Critères de qualité	89
L'emphase dans le processus.....	91
Les erreurs dans les bases de données.....	92
Erreurs d'indexation.....	92
Erreurs d'orthographe et de dactylographie	94
Erreurs d'omission	95
Erreurs de duplication.....	96
Infométrie, scientométrie et des indicateurs de R&D.....	102
Les différents types d'indicateurs.....	103
L'expérience française de l'OST	105
Indicateurs de R&D au Brésil	107
Les polémiques sur les indicateurs	110
La création de systèmes d'intelligence à l'Embrapa.....	115
L'agronoçoce et le futur.....	115
Les objectifs stratégiques de l'Embrapa	118
La politique de communication de l'Embrapa.....	119
Intelligence compétitive à l'Embrapa	121
Le projet "Développement de bases de données"	125

Audit de la collection documentaire et analyse de la production scientifique.....	127
L'Ainfo et son histoire	127
L'audit de la collection documentaire.....	132
Une solution bibliométrique pour la création d'indicateurs	134
Le traitement de la collection documentaire et de la production scientifique.....	141
Méthodologie	142
Auteur.....	144
Titre	150
Année de publication.....	152
Doublons	155
Exécution.....	156
Résultats du nettoyage.....	157
Limites du travail.....	161
Les limites du traitement automatique.....	161
Les limites des bases collection documentaire et production scientifique.....	174
Les limites du logiciel Ainfo	175
L'analyse de la production scientifique de l'Embrapa	177
Méthodologie	177
Résultats	182
Comparaison des bases Collection documentaire et Production scientifique.....	184
Comparaison des bases Pascal et WOS	186
La base Pascal.....	189
Groupes de collaboration	189
L'analyse des institutions	194
Le groupe Embrapa	195
L'évolution des partenariats	197
La vision des partenariats par centres de recherche	199
L'étude des revues scientifiques	200
L'étude des langues.....	201
La base WOS	203
Le groupe Embrapa	203
Les caractéristiques des distributions par régions de l'Embrapa	205
Les partenaires du groupe Brésil	206

Les caractéristiques régionales de la coopération au Brésil.....	207
Les partenariats dans le monde.....	208
Les groupes EUA/Canada et Europe.....	210
Les groupes Afrique/Amérique Latine/Asie et Système ONU.....	212
L'étude des revues scientifiques	214
L'étude des langues.....	217
Conclusion.....	219
Analyse stratégique des compétences de l'Embrapa.....	220
Le développement des bases de compétences.....	220
En approfondissant l'analyse des compétences.....	231
La croissance des sorties en 2005	240
Conclusion.....	247
Analyse de la performance de l'Embrapa dans les médias.....	249
Le système de Clipping électronique de l'Embrapa.....	249
Objectifs	250
Methodologie	251
La base Clipping	251
Véhicules	254
Centres de recherche.....	254
États (départements).....	255
Année de publication.....	255
Nombre de pages	255
Matrices	255
Segment véhicules	256
Résultats	257
Genre journalistique, traitement graphique et type de présence.....	260
L'évolution des dépêches pas mois.....	262
Les politiques éditoriales des véhicules	262
Les polarisations des éditoriaux de deux journaux nationaux, une approximation....	269
Les caractéristiques des couvertures.....	270
Une étude du Rio de Janeiro	271
Les politiques éditoriales par segment de véhicules.....	275
Les efforts éditoriaux des centres de recherche.....	276

L'analyse des thèmes	282
Conclusion.....	285
Conclusion Générale	288
Audit et analyse de la production scientifique	288
Les compétences de R&D	289
Le clipping.....	289
Bibliographie	293
Annexes	311
I. Les dix éditoriaux de l' "O Estado de S. Paulo"	311
II. Les cinq editoriaux de l' "O Globo"	324

INDEX DES FIGURES

Figure 1 : Le processus de génération de l'intelligence (Faria et Quoniam, 2002, apud. AFNOR, 1998).....	65
Figure 2 : Le processus de construction d'indicateurs à l' OST (Barré, 1997, p.2).....	106
Figure 3 : Le cycle du traitement automatique de l'information (Penteado et al., 2002, p.8)	137
Figure 4 : Production scientifique de l'Embrapa par année de publication (Penteado et al., 2002, p.10).....	139
Figure 5 : Production scientifique de l'Embrapa par chercheur (Penteado et al., 2002, p.11)	139
Figure 6 : Production scientifique de l'Embrapa par centre (Penteado et al., 2002, p.12) ...	140
Figure 7 : Carte des groupes d'auteurs, 1998.....	190
Figure 8 : Carte des groupes d'auteurs, 1999.....	191
Figure 9 : Carte des groupes d'auteurs, 2000.....	192
Figure 10 : Carte des groupes d'auteurs, 2001.....	193
Figure 11 : Carte des groupes d'auteurs, 2002.....	194
Figure 12 : Répartition des doctorats passés à l'étranger (Penteado et al., 2003a, p.3)	223
Figure 13 : Pyramide des âges de l'Embrapa (Penteado et al., 2003a, p.3).....	223
Figure 14 : Personnels de l'Embrapa, répartition des thèses de Doctorat par an (Penteado et al., 2003a, p.3)	224
Figure 15: Étude de la consanguinité des chercheurs de l'Embrapa Cerrados (Boutin, 2001, p.7)	225
Figure 16 : Réseau des compétences économiques par produits de l'Embrapa (Penteado et al., 2003a, p.4)	228
Figure 17 : Détail du réseau précédent (Penteado et al., 2003a, p.5).....	229
Figure 18 : Distribution des chercheurs par cohorte (Penteado et al., 2005a, p.24)	233
Figure 19 : Compétences des plus jeunes (Penteado et al., 2005a, p.25)	234
Figure 20 : Compétences en risque (Penteado et al., 2005a, p.26)	235
Figure 21 : Où sont les plus jeunes (Penteado et al., 2005a, p.27)	236
Figure 22 : Où sont les plus vieux (Penteado et al., 2005a, p.28).....	237
Figure 23 : Évolution des compétences par secteur de recherche Cohortes 1 et 4 (Penteado et al., 2005a, p.33).....	238

Figure 24 : Évolution des compétences par secteur de recherche Cohortes 2 et 4 (Penteado et al., 2005a, p.34).....	239
Figure 25 : Compétences perdues (Penteado et al., 2005a, p.35)	240
Figure 26 : 303 Chercheurs sortants depuis 1996	240
Figure 27 : Page du Clipping Électronique avant digitalisation.....	252
Figure 28 : Entrée de la page digitalisée dans le Lotus Notes.....	253

INDEX DES TABLEAUX

Tableau 1 : Sources formelles d'information (Coelho, 1999, apud. Penteado, 2001, p.28).....	63
Tableau 2 : Sources informelles d'information (Coelho, 1999, apud. Penteado, 2001, p.28)..	63
Tableau 3 : Variables pour comparer des références à l'UCAL (Hickey et Rypka, 1979, p.134)	98
Tableau 4 : Variables pour l'analyse des logiciels de bibliothèques (Corte et d'autres, 2002)	101
Tableau 5 : Base Collection documentaire ; Nombre de références indexées 1996/2005.....	128
Tableau 6 : Base Production scientifique ; Références valides, avril 2005.....	131
Tableau 7 : Le registre Collection documentaire original	142
Tableau 8 : Résultat du reformatage du champ "année de publication"	154
Tableau 9 : Le registre Collection documentaire traité.....	157
Tableau 10 : Auteurs, base Collection documentaire originale et traitée	159
Tableau 11 : Les 20 auteurs le plus "bénéficiés" par le reformatage	160
Tableau 12 : Petite échantillon des 152.021 auteurs et 284.950 titres avec fréquence un	161
Tableau 13 : Trois références similaires (ABREU, C. A.)	161
Tableau 14 : Trois titres similaires	162
Tableau 15 : Autres trois titres similaires du tableau 12.....	163
Tableau 16 : Quelle-est l'année correcte-1?.....	164
Tableau 17 : Quelle-est l'année correcte-2?.....	165
Tableau 18 : Quel registre conserver-1?	166
Tableau 19 : Quel registre conserver-2?	166
Tableau 20 : Références similaires mais pas des doublons	167
Tableau 21 : Références valides par pourcentage de la Production scientifique, avril 2005 .	173
Tableau 22 : Le registre WOS original.....	179
Tableau 23 : Le registre WOS traité	180
Tableau 24 : Le registre Pascal original.....	180
Tableau 25 : Le registre Pascal traité	181
Tableau 26 : Production scientifique de l'Embrapa de 1998 à 2002, par centre	183
Tableau 27 : Auteurs, bases Production scientifique et Collection documentaire 1998/2002	185
Tableau 28 : Auteurs, bases WOS et Pascal, 1998/2002.....	187
Tableau 29 : Base Pascal ; Groupe Embrapa.....	196

Tableau 30 : Base Pascal ; Participation en articles par institution et année.....	198
Tableau 31 : Base Pascal ; Participation en articles par région géographique et année	199
Tableau 32 : Base Pascal ; Partenaires de l'Embrapa Ressources Génétiques et Biotechnologie (Cenargen), à gauche et de l'Embrapa Instrumentation Agricole (CNPDI), à droite ..	200
Tableau 33 : Base Pascal ; Revues scientifiques préférées à l'Embrapa Ressources Génétiques et Biotechnologie, à gauche et à l'Embrapa Instrumentation Agricole, à droite.....	201
Tableau 34 : Base Pascal ; Langues des articles	202
Tableau 35 : Base WOS ; Groupe Embrapa.....	204
Tableau 36 : Base WOS ; Participations en articles de l'Embrapa par région et année.....	205
Tableau 37 : Base WOS ; Partenaires Embrapa de l'UFV, l'UNB, l'UNESP et l'UFMG.....	207
Tableau 39 : Base WOS; Partenaires Embrapa au Brésil par région et année	208
Tableau 38 : Base WOS ; Les centres de l'Embrapa les plus internationaux.....	209
Tableau 40 : Base WOS ; Les 13 institutions de l'EUA/Canada avec plus de participations en articles par année.....	211
Tableau 41 : Base WOS ; Les 13 premiers partenaires de l'Embrapa en Europe par année ..	212
Tableau 42 : Base WOS ; Les dix premiers partenaires de l'Embrapa en Afrique/Amérique Latine/Asie par année	213
Tableau 43 : Base WOS ; Les partenaires de l'Embrapa en le Système ONU par année.....	213
Tableau 44 : Base WOS ; Participation en articles par revue scientifique	214
Tableau 45 : Base WOS ; Les trois revues les plus préférées par les centres Embrapa	216
Tableau 46 : Base WOS ; Groupe Embrapa, participation en articles par centre et par langue	218
Tableau 47 : Structure originale de la base Compétences (Penteado et al., 2003a, p.2)	221
Tableau 48 : Les champs documentaires (Penteado et al., 2003a, p.2)	221
Tableau 49 : Structure modifiée de la base Compétences (Penteado et al., 2003a, p.3)	222
Tableau 50 : Exemple de tableau multi-compétences - Embrapa (Penteado et al., 2003a, p.5)	230
Tableau 51 : Compétences perdues par secteurs de recherche.....	242
Tableau 52 : Secteurs de recherche des chercheurs avec 55 ans ou plus.....	244
Tableau 53 : Lignes de recherche avec plus de chercheurs seniors	246
Tableau 54 : Les nouvelles variables de la base Clipping	251
Tableau 55 : La base Clipping reformatée.....	254
Tableau 56 : Les matrices générées avec la base Clipping.....	255

Tableau 57 : Les matrices des segments de la variable véhicules.....	257
Tableau 58 : Base Clipping ; Statistique des champs vides	258
Tableau 59 : Base Clipping ; Les 50 véhicules ayant publié le plus de nouvelles.....	259
Tableau 60 : Base Clipping ; Genre journalistique par année de publication	260
Tableau 61 : Base Clipping ; Traitement graphique par année de publication	261
Tableau 62 : Base Clipping ; Présence éditoriale par année de publication.....	261
Tableau 63 : Base Clipping ; Evolution des nouvelles par mois.....	262
Tableau 64 : Véhicules concédant le meilleur traitement graphique à l'Embrapa.....	264
Tableau 65 : Véhicules publiant le plus de dépêches simples sur l'Embrapa.....	266
Tableau 66 : Véhicules ayant cité le plus à l'Embrapa dans leur première page	267
Tableau 67 : Véhicules publiant le plus d'éditoriaux sur l'Embrapa	268
Tableau 68 : Véhicules publiant le plus de reportages et notes d'opinion sur l'Embrapa.....	269
Tableau 69 : Centres de recherche de l'Embrapa par état et véhicules par état	271
Tableau 70 : Véhicules de l'état du Rio de Janeiro ayant publié des dépêches des centres de recherche de l'Embrapa du Rio	272
Tableau 71 : Les 26 véhicules ayant publié jusqu'à trois nouvelles des centres du Rio.....	274
Tableau 72 : Centres par région et les journaux nationaux	275
Tableau 73 : Centres par région et les principaux journaux.....	276
Tableau 74 : Centres par nombre de nouvelles et année.....	277
Tableau 75 : Véhicules par Centre ; Siège Embrapa (DF), Embrapa Soja (PR), Embrapa Blé (RS), Embrapa Bovins de Lait (MG), Embrapa Élevage du Bétail Sud-Est (SP) et Embrapa Sols (RJ), vision 1	279
Tableau 76 : Évolution de la couverture par véhicule par centre ; Ex. Embrapa Soja	280
Tableau 77 : Centres par véhicule, "O Estado de S. Paulo", vision 2.....	281
Tableau 78 : Centres par véhicule, "Gazeta Mercantil", vision 2	282
Tableau 79 : Véhicules par mots-clés (transgéniques)	283
Tableau 80 : Véhicules par mots-clés (transgéniques) par nombre de nouvelles	284
Tableau 81 : Véhicules par mots-clés (transgéniques), importance relative des nouvelles ...	284
Tableau 82 : Véhicules par état et mots-clés	285

INDEX DES ABRÉVIATIONS

ABRÉVIATIONS EMBRAPA

Embrapa	Entreprise Brésilienne de Recherche Agronomique
EMBRAPA	Empresa Brasileira de Pesquisa Agropecuária
SIÈGE	Siège de l'Embrapa
UD	Unité décentralisée de l'Embrapa (centre de recherche)
CAFE	Embrapa Café
Genargen	Embrapa Ressources Génétiques et Biotechnologie
CENARGEN	Embrapa Recursos Genéticos e Biotecnologia
CNPA	Embrapa Cotton
CNPA	Embrapa Algodão
CNPAB	Embrapa Agrobiologie
CNPAB	Embrapa Agrobiologia
CNPAF	Embrapa Riz et Haricot
CNPAF	Embrapa Arroz e Feijão
CNPAT	Embrapa Agro-industrie Tropicale
CNPAT	Embrapa Agroindústria Tropical
CNPC	Embrapa Caprins
CNPC	Embrapa Caprinos
CNPDIA	Embrapa Instrumentation Agricole
CNPDIA	Embrapa Instrumentação Agropecuária
CNPF	Embrapa Forêts
CNPF	Embrapa Florestas
CNPGC	Embrapa Bovins de Viande
CNPGC	Embrapa Gado de Corte
CNPGL	Embrapa Bovins de Lait
CNPGL	Embrapa Gado de Leite
CNPH	Embrapa Horticulture
CNPH	Embrapa Hortaliças
CNPM	Embrapa Surveillance par Satellite
CNPM	Embrapa Monitoramento por Satélite
CNPMA	Embrapa Environnement

CNPMA	Embrapa Meio Ambiente
CNPMPF	Embrapa Manioc et Arboriculture Fruitière Tropicale
CNPMPF	Embrapa Mandioca e Fruticultura Tropical
CNPMS	Embrapa Maïs et Sorgho
CNPMS	Embrapa Milho e Sorgo
CNPS	Embrapa Sols
CNPS	Embrapa Solos
CNPISA	Embrapa Porcs et Volailles
CNPISA	Embrapa Suínos e Aves
CNPISO	Embrapa Soja
CNPIT	Embrapa Blé
CNPIT	Embrapa Trigo
CNPITIA	Embrapa Informatique Agricole
CNPITIA	Embrapa Informática Agropecuária
CNPUIV	Embrapa Raisin et Vin
CNPUIV	Embrapa Uva e Vinho
CPAA	Embrapa Amazonie Occidentale
CPAA	Embrapa Amazônia Ocidental
CPAC	Embrapa Cerrados
CPACT	Embrapa Climat Temperé
CPACT	Embrapa Clima Temperado
CPAF-AC	Embrapa Acre
CPAF-AP	Embrapa Amapa
CPAF-RO	Embrapa Rondonia
CPAF-RR	Embrapa Roraima
CPAMN	Embrapa Moyen-Nord
CPAMN	Embrapa Meio-Norte
CPAO	Embrapa Agriculture et Élevage Ouest
CPAO	Embrapa Agropecuária Oeste
CPAP	Embrapa Pantanal
CPATC	Embrapa Plateaux Côtiers
CPATC	Embrapa Tabuleiros Costeiros

CPATSA	Embrapa Semi-Aride
CPATSA	Embrapa Semi-Árido
CPATU	Embrapa Amazonie Orientale
CPATU	Embrapa Amazônia Oriental
CPPSE	Embrapa Élevage du Bétail Sud-Est
CPPSE	Embrapa Pecuária Sudeste
CPPSUL	Embrapa Élevage du Bétail Sud
CPPSUL	Embrapa Pecuária Sul
CTAA	Embrapa Agro-industrie d'Aliments
CTAA	Embrapa Agroindústria de Alimentos
SCT	Embrapa Information Technologique
SCT	Embrapa Informação Tecnológica
SNT	Embrapa Transfert de Technologie
SNT	Embrapa Transferência de Tecnologia

ABRÉVIATIONS DES ÉTATS DU BRÉSIL, PAR RÉGION

RÉGION NORD

AC	Acre
AM	Amazonas
AP	Amapa
PA	Para
RO	Rondonia
RR	Roraima
TO	Tocantins

REGION NORD-EST

AL	Alagoas
BA	Bahia
CE	Ceara
MA	Maranhao
PB	Paraiba
PE	Pernambuco
PI	Piaui

RN **Rio Grande do Norte**

SE **Sergipe**

REGION CENTRE-OUEST

DF **Distrito Federal**

DF District Fédéral

GO **Goias**

MS **Mato Grosso do Sul**

MT **Mato Grosso**

REGION SUD-EST

ES **Espirito Santo**

MG **Minas Gerais**

RJ **Rio de Janeiro**

SP **Sao Paulo**

REGION SUD

PR **Parana**

RS **Rio Grande do Sul**

SC **Santa Catarina**

AUTRES ABRÉVIATIONS

AAFCAN **Agriculture and Agri-Food Canada**

ABNT Agence Brésilienne de Normes Techniques

ACS Bureau de Communication Sociale de l'Embrapa

AFNOR Association française de normalisation

AHCI **Arts & Humanities Citation Index**

AHMADBUN Université d'Ahmadu Bello - Nigeria

AIENERGIE Agence Internationale d'Energie

ARS-USDA **Agricultural Research Service**

ASCII **American Standard Code for Information Interchange**

BDPA Base de données de la recherche agronomique de l'Embrapa

BiblioWOS Web of Science Bibliométrique

CAB **Commonwealth Agricultural Bureaux**

CE Commission Européenne

CÉDÉROM	Disques compacts pressés
CENA	Centre d'Énergie Nucleaire dans l'Agriculture
CIA	Central Intelligence Agency
CIAT	Centre International d'Agriculture Tropicale
CIFOR	Centre International de Recherche des Forêts
CIQM	Centre for Information Quality Management
CIRAD	Centre de Coopération Internationale en Recherche Agronomique pour le Développement
CFRNZ	Crop and Food Research Institute - Nouvelle Zélande
CNPq	Conselho Nacional de Desenvolvimento Científico e Tecnológico [Conseil National de Développement Scientifique et Technologique] du Brésil
CNRS	Centre National de la Recherche Scientifique
CP-Ainfo	Commission permanente pour l'Ainfo de l'Embrapa
CRRM	Centre de Recherches Retrospectives de Marseille - Université d'Aix Marseille III
CSIRO	Commonwealth Scientific and Industrial Research Organisation
CSU	Université de l'État du Colorado
CSV	Comma Separated Value File
DEA	Diplôme d'Études Approfondies
DGP	Département de Gestion de Personnes de l'Embrapa
DM	Datamining
DPD	Département de Recherche et Développement de l'Embrapa
DW	Datawarehouse
EQUIP	European Quality in Information Programme
ESALQ	Escola Superior de Agricultura Luiz de Queiroz
ETC	Extraction, Transformation et Charge
ETL	Extraction, Transformation and Loading
EUA	États-Unis
Fapesp	Fundação de Amparo à Pesquisa do Estado de Sao Paulo [Fondation de l'Aide à la Recherche de l'Etat de Sao Paulo]

Fiocruz	Fundação Oswaldo Cruz
FT-SEM	FT-Sementes
GSF	Institute of Soil Ecology , Neuherberg
GT-Refer	Groupe de travail de référence bibliographique de l'Embrapa
HTML	HyperText Markup Language
IAC-APTA	Institut Agronomique de Campinas
IACR-UK	Rothamsted Research
IBAMA	Institut Brésilien de l'Environnement et des Ressources Naturelles Renouvelables
IBGE	Institut Brésilien de Géographie et Statistique
IBICT	Instituto Brasileiro de Informação Científica e Tecnológica [Institut Brésilien d'Information Scientifique et Technologique]
IBM	International Business Machines
IC	Intelligence compétitive
ICRISAT	Institut International de Recherche Agronomique du Tropic Semi-Aride
ICT	Intelligence compétitive technique
ICV	Indice des conditions de vie
INIST	Institut de l'Information Scientifique et Technique
INRA	Institut Nationale de Recherche Agronomique
INSPEC	Physics Abstracts, Electrical & Electronics Abstracts, Computer & Control Abstracts, and Business Automation
INTA-ARG	Institut Nationale de Technologie Agricole - Argentine
ISBN	International Standard Book Number
ISI	Institute for Scientific Information
ISSN	International Standard Serial Number
LABEX	Laboratoire Virtuel à l'Extérieur de l'Embrapa
LEPONT	Laboratoire d'Etude pour l'Ouverture aux Nouvelles Technologies - Université du Sud, Toulon-Var
MAHATGANDHIUNI	Université Mahatma Gandhi - Inde
MCM	Moyens de communication de masse
MCT	Ministère de la Science et Technologie du Brésil

MGE	Modèle de Gestion Stratégique de l'Embrapa
MONSAN	Monsanto Corporation
MPO	Ministério do Planejamento, Orçamento e Gestão [Ministère du Plan, Budget et Gestion] du Brésil
MSIRI-MAURIT	Mauritius Sugar Industry Research Institute - Maurice
MSU	Université de l'État du Michigan
NARCWR-JAP	Centre National de Recherche Agronomique de la Région Ouest - Japon
NBR	Norme Brésilienne
NCSU	Université de l'État de North Caroline
NEC	Nippon Electric Company
NIST	National Institute of Standards and Technology
NIT Materiais	Noyau d'Informations Technologiques en Matériels de l'Université Fédérale de Sao Carlos
NORTHERNARIZUN	Université de Northern Arizona
NSF	National Science Foundation
NTC	Nouvelles technologies de la communication
OBAFEMIAWOU	Université d'Obafemi Awolowo - Nigeria
OCDE	Organisation pour la Coopération et le Développement Économique
OEs	Objectifs stratégiques de l'Embrapa
ONPs	Organisations qui ne visent pas à des profits
OST	Observatoire des Sciences et des Techniques
PAT	Plan annuel de travail de l'Embrapa
PDE	Plan Directeur de l'Embrapa
PDF	Portable Document File
PEA	Population économiquement active
PIB	Produit Intérieur Brut
PUC-EQUAD	Pontificia Université Catholique de l'Equador
QPAP	Programme de qualité et participation dans l'administration publique
R&D	Recherche et Développement

RICYT	Red Iberoamericana de Indicadores de Ciencia y Tecnologia
S&T	Science et Technologie
SAAD	Système de planification, d'accompagnement et d'évaluation de résultats du travail individuel de l'Embrapa
SAINSBURYLAB-UK	John Innes Centre, Norwich
SAPRE	Système d'évaluation et de récompense par des résultats de l'Embrapa
SAU	Système d'évaluation des unités de l'Embrapa
SCIE	Science Citation Index Expanded
SCOUG	Southern California Users Group
SEG	Système Embrapa de gestion
SEP	Système Embrapa de planification
SIRH	Système d'information des ressources humains de l'Embrapa
SISPAT	Système plan annuel de travail de l'Embrapa
SNECMA	Société nationale d'étude et de construction de moteurs d'aviation
SNPA	Système National de Recherche Agronomique du Brésil
SPD	Surintendance de Recherche et Développement de l'Embrapa
SRDP	Standart Reference Data Program
SSCI	Social Science Citation Index
UARIZ	Université de l'Arizona
UATHENS	Université d'Athènes
UBAYREU	Université de Bayreuth
UBAYREUT	Université de Bayreuth
UBRITCOL	Université de British Columbia
UBRITCOLUMBIA	Université de British Columbia
UCAL	Université de Californie
UCATLOUVAIN	Université Catholique de Louvain
UCB	Université Catholique de Brasilia
UCORNEL	Université de Cornell
UEMARINGA	Université Étatique de Maringa
UESC	Université Étatique de Santa Catarina
UFLA	Université Fédérale de Lavras

UFLOIDA	Université de la Florida
UFMG	Université Fédérale de Minas Gerais
UFOP	Université Fédérale d'Ouro Preto
UFPE	Université Fédérale de Pernambuco
UFPR	Université Fédérale du Paraná
UFRGS	Université Fédérale du Rio Grande du Sud
UFRJ	Université Fédérale du Rio de Janeiro
UFRRJ	Université Fédérale Rurale du Rio de Janeiro
UFSCAR	Université Fédérale de Sao Carlos
UFSM	Université Fédérale de Santa Maria
UFV	Université Fédérale de Viçosa
UGEORGIA	Université de Georgia
UGOETTINGEN	Université de Göttingen
UHAMB	Université de Hamburg
UHAMBURG	Université de Hamburg
UHOKKAIDO	Université de Hokkaido
ULEEDS	Université de Leeds
ULINCLN-NZ	Université de Lincoln, Nouvelle Zélande
ULONDON	Université de Londres
UNAC-COL	Université Nationale de Colombie
UNACCUZ-PER	Université Nationale de San Antonio Abad Cusco, Perou
UNACNE-ARG	Université Nationale du Nordeste, Argentine
UNACSUR-ARG	Université Nationale du Sud, Argentine
UNAM	Université Nationale Autonome du Mexique
UNASSUMP-PAR	Université Nationale d'Asunción, Paraguay
UNB	Université de Brasilia
UNEAFIT-COL	Université EAFIT, Medellin, Colombie
UNESCO	Organisation des Nations Unies pour l'Éducation, la Science et la Culture
Unesp	Université Paulista
UNESP	Université Paulista
Unicamp	Université de Campinas

UPENNS	Université de Pennsylvania
UPURDUE	Université de Purdue
UQUEENSLAND	Université de Queensland
UREADING	Université de Reading
USBC	Universal standart bibliographic code
USDA-ARS	Agricultural Research Service
USP	Université de Sao Paulo
USP-SC	Université de Sao Paulo - Sao Carlos
UTAHSU	Université de l'État d'Utah
UTOKYO	Université de Tokyo
UVALENCIA	Université de Valencia
UWAGENINGEN	Université de Wageningen
UWISMAD	Université de Wisconsin-Madison
UWISCAD	Université de Wisconsin-Madison
WOS	Web of Science

INTRODUCTION

Introduction

Cette thèse est un exercice d'union entre les domaines de la Science de l'Information et de la Communication. Elle est aussi le couronnement et l'intégration du travail et la réflexion d'une vie.

La dissertation défendue à l'Université de Floride, en 1996, pour obtenir le titre de Master of Arts, sur "Les effets de Fonctions et Modèles de Relations Publiques dans des Organisations Brésiliennes Compromises avec la Qualité" a consolidé la notion selon laquelle

l'environnement est devenu chaque fois plus important pour les organisations et la communication est et doit être un acteur et instrument de cette interaction entre l'Entreprise et l'Environnement. Cruciale aussi la maxime de Grunig et Repper (1992) selon laquelle les relations publiques sont un instrument stratégique pour interconnecter les organisations et leurs environnements et cette fonction doit être administrée stratégiquement, en développant des relations avec des associés stratégiques dans les environnements interne et externe, de façon à élargir la capacité de l'organisation d'accomplir sa mission (Penteado, 1996, p.25).

Le mémoire de Diplôme d'Études Approfondies (DEA) soumis à l'Université du Sud, Toulon-Var en 2001, sur des "Systèmes d'Intelligence Compétitive pour la Gestion Stratégique dans une Institution de Recherche et Développement" a défini des objectifs, un modèle conceptuel et une proposition pour la création de systèmes d'intelligence pour la communication de l'Embrapa, dans la recherche d'améliorer les relations et la connaissance mutuelle entre l'Embrapa et ses publics.

Le présent travail reprend la discussion de ces deux études antérieures et va plus loin, en pénétrant dans le processus d'analyse et de compréhension de l'environnement et de ces différentes variables, en créant des possibilités d'ajouter, quand elles sont nécessaires, des informations stratégiques pour l'organisation, de manière à améliorer le processus de décisions.

Pour Dou (2001, 2003) les deux compétences cruciales d'une organisation moderne sont la capacité d'acquérir de la connaissance et l'habileté de traiter des informations. Au contraire de ce que croit une majorité, la connaissance et l'information pour l'action, l'intelligence, "ne sont pas trouvées". L'intelligence et la connaissance "doivent être créés". Pour citer Levet (2001,

p.38), "ce n'est plus l'accès à l'information qui est le ressort de la croissance et de l'emploi mais l'aptitude des acteurs en transformer, comprendre, interpréter et utiliser l'information". Dutheil (1991) a enregistré cette tendance dans son "Rapport de fin de contrat SGDN" où il affirme que le pouvoir n'appartient plus à celui qui détient l'information mais à celui qui sait la traiter.

Les organisations et leurs gérants, dans cette période de transition entre la deuxième et la troisième vague (Toffler, 1980), doivent être conscients de la nécessité de prendre des mesures administratives pour amener à une effective gestion de l'information et de la connaissance et à une capacité de traiter des informations pour l'action et pour améliorer les décisions. La capacité d'utiliser des informations déjà disponibles dans l'organisation et de les recombinaison pour générer d'autres informations plus agrégées et la capacité à rechercher des informations stratégiques au dehors de l'organisation et de les intégrer au processus stratégique nécessitent d'être nourries et cultivées.

Pourtant, l'arme décisive de ceux qui veulent être des acteurs dans la nouvelle société de l'information et de la communication est la capacité d'analyser, en temps réel, le maximum d'informations possibles et d'établir, à partir de cette masse, une quantité significative d'informations pour l'action : Des intelligences économiques, financières, juridiques, diplomatiques, culturelles, sociales, scientifiques et politiques. Comme enseigne Dou (1999), percevoir immédiatement un signal faible, un petit détail, peut amener à des grands avantages compétitifs.

Le Plan de Thèse

Cette thèse traite de ce processus de création de systèmes d'intelligence, ses résultats opérationnels et stratégiques, du traitement de l'information et de son application dans une organisation publique de recherche et développement. Nous avons commencé par une série de définitions et synthèses de concepts pour former une base théorique : Société de l'information et de la communication, communication et nouvelles technologies, données, information et systèmes, connaissance, connaissance de l'environnement, communication, intelligence, veille et leur dimension économique et sociale.

Nous traitons ensuite de la nécessité d'information des organisations modernes, du cycle de l'intelligence et des étapes de préparation, analyse et synthèse d'informations et de données,

soit sur la forme textuelle soit sur la forme chiffrée. Nous avons cherché à définir ce qui serait la qualité dans une base de données et d'avoir accès à l'influence que telles normes ont dans la génération d'indicateurs, à partir de ces bases. Nous avons discuté les expériences française, de l'Observatoire des Sciences et des Techniques (OST) et brésilienne, de la **Fundação de Amparo à Pesquisa do Estado de Sao Paulo** [Fondation de l'Aide à la Recherche de l'Etat de Sao Paulo] (Fapesp), dans la construction d'indicateurs et les polémiques sur l'utilisation et la génération de tels indices.

L'incorporation de ces méthodologies dans les processus administratifs de l'Entreprise Brésilienne de Recherche Agronomique (Embrapa) a eu comme résultats une initiative de scientométrie pour l'audit de la collection documentaire et l'analyse de la production scientifique de l'Embrapa, une initiative pour l'analyse des compétences de recherche et développement de l'Embrapa et une autre de médiamétrie pour l'analyse de la performance de l'Embrapa dans les médias.

Pendant que les études de scientométrie apportent des contributions singulières, nous considérons que le plus innovateur est celui de médiamétrie parce qu'il contient des caractéristiques peu exploitées, soit dans le domaine de la Science de l'Information soit dans le domaine de la Science de la Communication. Il place sur des bases solides une méthodologie bien fondée et robuste d'analyse bibliométrique pour analyser les moyens de communication de masse. Ainsi il devient possible de générer des tendances de la couverture de la presse et de déterminer les préférences des divers véhicules, ainsi comme faire des études prospectives sur l'évolution historique d'un thème, d'un produit, d'un politicien ou d'une entreprise dans les médias et encore retirer, de cette analyse, des conclusions pour orienter et perfectionner une relation, en somme, pour créer des liens plus profonds et des meilleures relations, encore plus productives entre le sujet de recherche et les médias.

L'étude du journalisme dans sa pratique, aujourd'hui, dans le monde, les relations publiques et la communication d'entreprise, le marketing et aussi le segment des audit-conseil des médias ont à disposition, des outils et méthodologies qui permettent un saut de qualité et de productivité dans toutes les analyses médiatiques réalisées jusqu'à maintenant.

DÉFINITIONS ET SYNTHÈSES

Définitions et synthèses

Société de l'information et de la communication

La nouvelle société de l'information est le résultat, selon Tarapanoff (2001, p.36), de la convergence de nouveaux référentiels sociaux, économiques, technologiques et culturels :

- L'information est la principale matière-première, une composante comparable à l'énergie qui alimente un système ;
- La connaissance est utilisée ajoutant de la valeur à de produits et services ;
- La technologie constitue un élément vital pour le changement, spécialement l'emploi de technologie sur les réservoirs des collections d'information ;
- La rapidité, l'effectivité et la qualité constituent des facteurs décisifs de compétitivité.

Daniel Bell (1973, p.467) argumente que le savoir et la technologie sont à la base de la "société post-industrielle" et l'information devient, dans tous les niveaux de la vie sociale et du travail, une source de pouvoir. Néanmoins, Masson (2001) indique qu'en fonction de la quantité et de la qualité, d'être chaque fois plus technique et, parfois, un peu obscure, "l'information exige des connaissances chaque fois plus profondes" et "la principale difficulté est dans l'analyse et le traitement de l'information" (p.46).

Ce nouveau système économique basé en la connaissance et en l'information a permis une nouvelle forme d'organisation des facteurs de production. Selon Jacques Manardo (Norlain et La Spiere, 1999), "cette nouvelle économie se caractérise par l'accélération des mutations, l'augmentation du terrain de compétition, la complexité des relations, la croissance des échelles de production, la multiplicité des auteurs, l'augmentation des incertitudes et des opportunités" (p.9).

Une des caractéristiques essentielles de cette nouvelle ère de la connaissance, telle qu'elle a été définie par Lyle et McLeod (1993, p.2) est "l'accélération des processus de changement - des institutions, du gouvernement, des entreprises et de la culture".

Toffler (1980) souligne que

cette nouvelle civilisation, dans la mesure en qu'elle adresse un défi à l'ancien, ira revigorer les bureaucraties, réduire le rôle de l'État Nation et inciter des économies semi-autonomes dans un monde post impérialiste. Elle requiert des gouvernements plus simples, plus effectifs et encore plus démocratiques que quelques-uns de ces connus aujourd'hui. C'est la civilisation avec son propre et distincte vision du monde, sa manière propre de traiter le temps, la logique et la causalité (pp.10-11).

Toffler (1990) note aussi qu'"une structure de pouvoir radicalement différente est en train de se former. Ceci se passe dans tous les niveaux de la société humaine" (p.27). Quant à la révolution de la communication, expliquent Lyle et McLeod (1993), elle se manifeste par la "croissance continue de notre habileté d'échanger des informations plus rapidement et plus amplement" (p.5). C'est pour cela que Penteadó (1996) affirme que

les organisations doivent faire face à deux défis fondamentaux dans cette nouvelle ère. En première place, elles doivent changer la manière de faire ses affaires. En deuxième place, elles doivent changer la manière de penser son affaire (p.32).

Dans cette nouvelle ère, l'information est une marchandise et a une valeur. L'information, nous enseignent Lyle et McLeod (1993, p.3), peut donner un avantage compétitif à des individus et à des entreprises et peut être vendue si elle peut être utilisée par d'autres et est fournie de première main.

Communication et nouvelles technologies

Le modèle social

Cooley (1909) a été l'un des premiers à détecter l'émergence de la révolution de la communication. Il affirme que cette révolution est liée à l'expansion de la nature humaine, à l'expansion de son pouvoir d'exprimer soi-même dans des ensembles sociaux, d'organiser la société chaque fois plus sur les hautes facultés de l'être humain, l'intelligence et la sympathie au lieu d'autorité, caste et routine. Telle révolution, argumente Cooley, place la communication et les moyens de communication de masse (MCM) au centre des attentions de la Politique. Elle rend possible l'opinion publique et comme corollaire la démocratie laquelle, dans sa conception est l'opinion publique organisée. La communication dans le sens de Cooley élève la nature humaine rendant possible le contact et, d'une manière indirecte, favorise le développement de l'intelligence, le déclin des façons mécaniques et arbitraires d'organisation, bien comme le surgissement d'un type de société plus humaine. C'est un

mécanisme par lequel les relations humaines existent et se développent. Elle inclut tous les symboles de l'esprit à côté des moyens pour les transmettre par l'espace et les préserver dans le temps.

Selon Lyle et McLeod (1993)

la communication est l'un des processus cruciaux par lesquels les individus et la société réussissent à progresser. Les moyens de communication de masse sont des institutions qui facilitent ce processus (p.1) (...) peut-être, la promesse la plus cruciale de l'ère de l'information pour le publique, en général, est que les individus ont leur habileté de communiquer expansée. Ceci signifie plus de flexibilité pour partager des informations avec des tiers et pour sélectionner et apprécier une richesse de connaissances et d'entretien (p.5).

Lasswell (1948) a défini le modèle de communication de main unique. Il a décrit la meilleure façon de représenter les actes de communication : (a) Qui (b) Dit quoi (c) Dans quel canal (d) Pour qui (e) Avec quels effets. C'est pour lui un acte unidirectionnel. Le communicateur envoie un message pour une audience.

Le modèle psychologique

Dans cette même année, Krech et Crutchfield (1948) ont décelé les principes généraux de la perception et de la cognition humaine. Ainsi, la communication a gagné encore une dimension psychologique, un modèle "psychologique", en addition au modèle "social" de Cooley et Lasswell.

Le modèle psychologique de la communication indique qu'il y a deux déterminants pour la perception: (a) des facteurs structuraux qui dérivent seulement de la nature, des stimuli physiques ou des effets nerveux évoqués dans les systèmes nerveux des individus; Et (b) des facteurs fonctionnels qui dérivent d'abord des nécessités, attitudes, expériences et des mémoires des individus. De cette assertion, Krech et Crutchfield en tirent deux lois basiques: (a) chaque fois plus grande sera la valeur sociale d'un objet, plus il sera susceptible d'être organisé par des déterminants comportementaux; Et (b) chaque fois plus grande sera la nécessité individuelle d'un objet socialement valorisé, plus elle sera marquée par des déterminants de comportement.

Comme corollaire, l'information est perçue et interprétée en termes des nécessités du récepteur, de ses émotions, de sa personnalité et de ses propres conditions d'acquisition de connaissance. Des objets et des événements qui sont proches l'un de l'autre tendent à être

appréhendés comme des parties d'une structure commune. Les structures de cognition et de connaissance sont organisées, en première place, en termes de proximité et similarité et nos perceptions sont déterminées essentiellement par nos jugements de valeur, par nos nécessités, nos réactions émotionnelles et par des causes et des effets.

Le modèle mathématique

En 1949, Claude Shannon et Warren Weaver ont lancé leur *Mathematical Theory of Communication* [Théorie Mathématique de la Communication] définie par Severin et Tankard (1992) comme "le plus important et influent stimulateur pour le développement d'autres modèles et théories en communication" (p.38). Rabaça et Barbosa (1978) décrivent les "éléments essentiels de ce modèle" (pp.109-110).

une source d'information sélectionne, d'un ensemble de messages possibles, un message; Le transmetteur convertit le message en des signaux, selon un code déterminé et ces signaux sont transmis pour le destinataire par un canal approprié; Le récepteur, de son côté, décode ou démodule les signaux reçus, pour récupérer le message original. "Quand je parle avec quelqu'un, mon cerveau est la source d'information, le cerveau de l'autre est le destinataire; Mon système vocal est le transmetteur et son oreille est le récepteur" (Warren Weaver) (pp.109-110).

Severin et Tankard (1992) ajoutent que

d'autres contributions capitales sont les concepts de Shannon et Weaver d'un message composé d'entropie et redondance et du nécessaire équilibre entre les deux pour une communication efficace, en tant qu'opposition au bruit d'un canal. (...) plus il y a de bruit dans un canal, plus la redondance est nécessaire, ce qui réduit la relative entropie du message. (...) En utilisant la redondance pour superposer au bruit du canal, on réduit la quantité d'information que peut être transmise dans un temps déterminé (pp.38-39).

Severin et Tankard (1992) indiquent que la technologie de la communication digitale est un résultat direct de la contribution de Shannon et Weaver au point de vue que la théorie mathématique de la communication est devenue connue comme la "théorie de l'information". Selon Rabaça et Barbosa (1978) la raison est parce qu'elle peut être appliquée aux plus divers contextes, "biologique, psychologique, sociale, linguistique", d'entre autres (p.109).

Rapportent Severin et Tankard (1992) que Shannon et Weaver utilisent le "terme **communication** (...) 'dans un sens très large pour inclure toutes les procédures selon

lesquelles un esprit peut affecter un autre' (Weaver, 1949, p.95)" (p.40). Le terme **information** a aussi un sens spécial. Et continuent Severin et Tankard :

Selon la théorie de l'information, l'information est très similaire à l'entropie dans les sciences physiques : Une mesure du degré d'incertitude. L'entropie est l'incertitude ou la désorganisation d'une situation. Dans la théorie de l'information elle est associée avec une quantité de liberté de choix, quand le message est construit (p.41).

Ainsi, le plus un message est organisé, plus le degré d'incertitude et d'entropie sera faible. Étant donné la connaissance qu'une personne a d'une langue (qui a une forte entropie ; dans ce cas, les règles de grammaire et de syntaxe) elle peut corriger des erreurs graphiques ou aller jusqu'à remplir des mots manquants dans un texte (Severin et Tankard, 1992, p.41).

Le modèle de main double

Ces trois versants de la science de la communication ont été intégrés en 1954 par Schramm (1954, Révisée 1971). Il argumente que les conditions dans lesquelles le message est transmis, les relations sociales du récepteur avec le communicateur, sa perception des conséquences sociales d'accepter ou d'agir en fonction du message, les conditions dans lesquelles le message a été reçu, les habilités du récepteur et ses réponses innées ou appréhendées doivent être considérées pour qu'on puisse prévoir, sans aucune ambiguïté, les conséquences d'un acte de communication.

Schramm (1971) a défini le modèle de communication de main double comme un processus de deux sens où l'audience n'est pas passive mais est un partenaire. Dans la communication, mentionne Schramm, il y a trois éléments et deux façons d'action: (a) Les éléments sont le communicateur, le message et le récepteur. (b) Les actions sont ainsi décrites : Le communicateur construit le message pour obtenir les réponses souhaitées. Le récepteur sélectionne un message parmi les messages disponibles et parmi les contenus disponibles, l'interprète et l'utilise selon sa volonté. Les deux actions sont séparées et motivées séparément mais elles s'unissent par le message.

Rabaça et Barbosa (1978) rapportent que Menezes (1973) a signalé une grande similarité entre tous les modèles de communication modernes avec le "schéma triplet de la communication présenté par Aristote : 1. La personne qui parle; 2. Le discours qu'elle prononce ; Et 3. La personne qui écoute" (p.109).

Les fonctions de la communication

Pour Schramm la communication est quelque chose que les personnes font. Alors, pour comprendre les processus de la communication, il est nécessaire de comprendre comment les personnes se rapportent les unes avec les autres. Du point de vue du communicateur, la communication sert pour : Informer, enseigner, amuser et proposer ou persuader. Du point de vue du récepteur elle sert pour : Comprendre, apprendre, divertir et disposer ou décider.

Schramm (1971, p.659) a amplifié la vision des fonctions de la communication de Lasswell, en ajoutant une quatrième fonction, "l'entretenement". Selon Schramm, la communication facilite quatre fonctions basiques dans la société humaine :

- **Survie** : Les nouvelles alertent les individus et membres de groupes des opportunités et des dangers ;
- **Coordination de réponse** : Discussion, persuasion et propagande cherchent à influencer la façon dont les individus et les groupes répondent aux opportunités et au danger ;
- **Transfert entre générations** : La civilisation est rendue possible par notre habilité d'accumuler connaissance, en repassant de l'information et de la connaissance aux générations suivantes ;
- **Entretenement** : La diversion, à côté du repos et de la relaxation est essentielle pour réduire la tension accumulée et la fatigue des individus et des communautés. Le divertissement peut aussi fonctionner comme un moyen d'attraction et de maintien de l'attention pour les communications plus sérieuses (apud. Penteado, 2002, pp.340-341).

Pour Lyle et McLeod (1993), celles-là sont les quatre fonctions basiques pour la société moderne. Elles expliquent pourquoi la communication est si importante pour la survie et le progrès des individus et de la société.

Les NTC

Lyle et McLeod ont une vision particulière des nouvelles technologies de la communication (NTC), rapportée par Penteado (2002) :

(a) le semi-conducteur, (b) le satellite d'orbite géosynchrone, (c) les fibres optiques, (d) la vidéo de haute résolution et (e) l'ordinateur sont en train de se transformer dans un véritable paradigme des sociétés modernes en révolutionnant toutes les sphères de la vie sociale. Dans le monde du travail, par moyen de la

digitalisation des processus productifs et par son intégration en réseau. Dans l'univers de la création de produits pour les médias, en passant par l'industrie du divertissement, qui pointe pour des façons de loisirs chaque fois plus interactives et virtuelles (p.341).

Et Penteadó continue

L'ordinateur facilite la communication personnelle. Avec la miniaturisation et la capacité de traitement chaque fois plus grande, il a la possibilité d'exécuter des tâches répétitives avec une grande vitesse et exactitude. Considéré comme un instrument de communication, l'ordinateur est une technologie encore au début de son développement. Il prospère avec l'incorporation du multimédia, en intégrant des divers moyens de communication et en profitant pour enlever ce qu'il y a de meilleur en chacun. Sa tendance, comme dans le cas de l'électroménager, est d'intégrer, chaque fois plus, dans un seul ensemble, la télévision, le téléphone, le fac-similé et la vidéocassette (p.341).

Cette faculté d'intégration des moyens est aussi la caractéristique du réseau mondial d'ordinateurs, l'Internet, en ajoutant encore le facteur de l'interactivité. Penteadó (2002) indique que Graham (1997, p.xii) "compare le développement de l'Internet à la naissance d'un nouveau paradigme de culture et de communication" :

Le modèle du réseau mondial d'ordinateurs rend facile et bon marché la distribution et l'accès à n'importe quelle forme de données digitales par quiconque - entreprise ou consommateur - avec des énormes conséquences pour l'économie, la culture et la société (p.341).

Penteadó (2002) cite une phrase d'un essai de Jessica Helfand, "Electronic Typography", cueillie par David Siegel (1996, p.xv), auteur d'un grand succès de librairie sur la construction de sites sur l'Internet :

Ici est la plus grande contribution pour la technologie de la communication, un réseau mondial en liant près de 50 millions de personnes partout dans le monde ... (p.341).

La croissance de l'Internet dans le monde est explosive, de l'ordre de 80 à 100% par an. Alors ces 50 millions d'utilisateurs en 1995/1996 sont estimés aujourd'hui (2005/2006) en plus d'un milliard. Au Brésil, les utilisateurs de l'Internet étaient estimés à 25,1 millions, en 2005 (**E-Consulting Corp.**, 2005).

Les nouvelles technologies de la communication ont des impacts sur les modèles de communication en réduisant les distances et le temps et en augmentant notre habilité d'échanger des messages ou des significations plus vite. Cutlip et al. (1985, p.103) citent une

phrase de Toffler : "la communication n'est pas la cause de cette transition, elle est plutôt son catalyseur". La société de la connaissance qui est en train de surgir, argumente Toffler (1980) "a amené avec elle des nouveaux styles familiaux, des manières différentes de travailler, d'aimer et de vivre, une nouvelle économie, des nouveaux conflits politiques et avant tout, une nouvelle conscience" (p.9).

Différences entre masse et public

Penteado (2002, p.342) observe que

Freidson (1953) a défini des différences entre masse et public : La masse est hétérogène. Elle est composée d'individus qui ne se connaissent pas, sont séparés les uns des autres dans l'espace et ne peuvent pas interagir ou échanger des expériences. La masse n'a pas un commandement défini et aucune ou une organisation sociale diffuse. La méthode appropriée pour étudier la masse donne à chaque individu un poids égal, en les classifiant selon des attributs démographiques comme l'âge, le sexe, les données socio-économiques et l'éducation. Les réactions de chaque individu dans une masse, pourtant, suivent celles de ses attributs démographiques et de sa personnalité.

D'autre côté, continue Freidson (1953), le public est homogène. Ce sont des individus qui ont quelques choses en commun, sont affectés par le même problème ou sujet. Le sociologue Herbert Blummer, cité par Grunig et Hunt (1984, p.143-144), clarifie encore plus ce concept. Il indique que public est un groupe de personnes qui: (1) Sont en train d'affronter un même problème. (2) Sont divisés sur la façon d'affronter le problème. Et (3) s'engagent en la discussion sur le problème. Cette définition est très proche de celle du philosophe John Dewey. Pour lui, le public est un groupe de personnes qui : (1) Sont en train d'affronter un problème similaire. (2) Reconnaittent que le problème existe. (3) S'organisent pour faire quelque chose sur ce problème (Grunig et Hunt, 1984, p.143-144).

Une conclusion logique est que des informations, des programmes, des campagnes destinées aux masses sont plus efficaces lorsqu'elles sont divulgués parmi des moyens de communication de masse. De la même forme, des campagnes, des programmes et des informations qui visent un public déterminé obtiennent une plus grande efficacité quand elles utilisent des véhicules spécifiques qui atteignent le public visé ou cible (p.342).

Penteado (2002, pp.347-348) signale qu'

avec l'Internet, le processus relativement lent et désarticulé de formation de publics décrit par Freidson s'accélère et s'organise rapidement. Les masses peuvent dépasser leur caractéristique d'être 'séparés un de l'autre dans l'espace' et trouver des intérêts convergents dans le Réseau et former, instantanément, des publics motivés et bien plus important, comme en témoignent les infinies salles de

chat disponibles à la carte, les forums et les listes de discussion chaque fois plus spécifiques, en focalisant sur un sujet particulier (p.347).

Holtz (1999, p.21) rappelle que les paradigmes connus par les communicateurs il y a très long temps passent par des "changements irrévocables". En fonction des NTC, ils surpassent le format traditionnel de communication de "peu pour plusieurs" pour ce nouveau format de "plusieurs pour plusieurs". Les audiences ne sont plus limitées à son rôle traditionnel de "consommatrices d'information". Elles peuvent interagir entre elles et avec les organisations de son intérêt.

Ces publics instantanés et plus dynamiques exercent, à leur tour, un important rôle dans la fiscalisation et le contrôle sur ce qui est communiqué. Les consommateurs et clients peuvent s'organiser autour d'un thème commun, faire des campagnes en ligne. Ils peuvent encore faire l'éloge ou des critiques, en des sites Internet spécifiques, d'une organisation ou de ses produits. Des informations tronquées ou erronées peuvent se disséminer ou être disséminées rapidement (p.348).

Relations publiques et marketing

Penteado (2002, pp.1-13) rapporte que

*l'une des définitions les plus débattues de relations publiques a été collectée par Broom et Tucker (1989). Ils ont relaté le résultat d'une conférence appelée pour faire des distinctions entre relations publiques et **marketing**. Après une journée de discussions, les conférenciers ont accordé que relations publiques et marketing traitent tous les deux "des rapports d'une organisation et emploient des processus similaires" (p.39). Entre les conférenciers étaient des noms bien connus comme William Ehling, professeur et directeur du Département de Relations Publiques de l'Université de Syracuse et Philip Kotler, professeur de **marketing** international de la Kellogg Graduate School of Management de l'Université Northwestern. Ils se sont aussi accordé sur ce qui pouvait réellement séparer les deux fonctions : Le résultat de ce que chacune cherche à obtenir.*

Les Relations publiques ont été définies comme:

[Le] processus administratif ayant comme but d'obtenir et de maintenir des accords et des comportements positifs entre les groupes sociaux desquels une organisation dépend pour accomplir sa mission. Sa responsabilité fondamentale est de bâtir et maintenir un environnement favorable pour une organisation (Broom et Tucker, 1989, p.40).

*[Le **marketing** a été défini comme] le processus administratif ayant comme but d'attirer et de satisfaire des clients dans le long terme pour atteindre aux objectifs économiques d'une organisation. Sa responsabilité fondamentale est de bâtir et de maintenir un marché pour des produits et services d'une organisation (p.40).*

*[Ainsi,] Relations publiques et **marketing** peuvent être faites par les mêmes personnes ou par différentes personnes. Le danger est que "le succès de l'un peut être contrebalancé par l'échec de l'autre" (Broom et Tucker, 1989, p.40). Les deux nécessitent de l'attention. Les deux traitent de relations qui sont fondamentales pour la survie de l'organisation avec son environnement [les relations publiques] et avec son marché [le **marketing**].*

Les relations publiques contribuent à la construction et au maintien d'un environnement favorable à une organisation. Le **marketing** contribue à la construction et au maintien d'un marché pour les produits de l'organisation. Dans ce travail, le terme **relations publiques** est utilisé, ainsi comme en Grunig (1992a, p.4), comme des synonymes d'**administration de la communication**, de **communication organisationnelle** et aussi des termes **communication sociale** et **communication d'entreprise**.

Données, information et systèmes

Penteado (2002, pp.342-343) dit que

Reynolds (1992) définit donnée comme "une réunion d'événements non organisés, qui n'ont pas été traités et transformés en information". Pour lui, ce n'est pas toujours qui les données dont nous avons besoin qui sont disponibles dans une forme utilisable. "Les données peuvent ne pas être dans les façons adéquates ou être incomplètes, imprécises ou dépassées," (p.29), affirme-t-il.

Déjà l'information est "la donnée qui quelqu'un considère utile pour un propos immédiat. Les données doivent être traitées pour être transformées en information utilisable" (p.29).

Reynolds (1992) conclut que "fréquemment l'information d'une personne peut être les données d'autre" (p.30) et décrit les cinq caractéristiques clés qui déterminent la qualité des données :

- *Précision - être livré d'erreurs ;*
- *Intégrité - contenir toutes les données ;*
- *Relevance - pertinence quant à la décision d'être utile ;*
- *Actualité - pour baser des décisions dans un environnement instable ;*
- *Audibilité - habileté de vérifier la précision et l'intégrité des données (2002, pp.342-343).*

Penteado argumente (2002, p.343) citant Reynolds (1992, p.31), que

des organisations qui réussissent doivent réunir des données, les analyser pour créer l'information, disséminer l'information pour le public approprié et interpréter cette information. (...) L'habileté d'une organisation à réunir des données, les interpréter et agir rapidement est une caractéristique qui peut distinguer vainqueurs et vaincus dans un marché extrêmement compétitif (p.31).

Penteado (2002, p.343) introduit "une autre définition de Reynolds (1992, p.33), celle de système" :

Un groupe d'éléments inter relationnels organisés pour atteindre un objectif commun. Tous les éléments du système ont une relation logique ; Ils travaillent tous dans la direction de l'objectif du système, au lieu d'avoir leurs propres objectifs (p.33).

Connaissance et information

Penteado (2002, p.343) indique que Nonaka et Takeuchi (1997, p.63) ont décrit "les similitudes et différences entre information et connaissance".

La connaissance, au contraire de l'information, se réfère à des croyances et à des compromis. La connaissance est fonction d'une attitude, perspective ou intention spécifique. (...) La connaissance, au contraire de l'information, a des relations directes avec l'action. Elle est toujours connaissance "avec une fin quelconque". (...) La connaissance, comme l'information, a des relations avec le contexte et elle est relationnelle.

Les auteurs japonais citent Bateson (1979), continue Penteado (2002, p.343), "l'information est composée de différences qui font la différence" (p.5) dès qu'elle proportionne nouvelles orientations pour interpréter des événements ou des objets. Ils font référence aussi à Machlup (1983) :

L'information rend visibles des significations avant invisibles ou jette la lumière sur des connexions inattendues. Elle est donc un milieu ou un matériel nécessaire pour extraire et construire la connaissance. Elle affecte la connaissance en l'ajoutant quelque chose ou en la restructurant (p.63).

C'est ainsi, continue Penteado (2002, p.343) que l'information, pour Nonaka et Takeuchi (1997), peut être encadrée de deux façons : Par son volume (par exemple : L'information syntactique) et par sa signification (exemple : L'information sémantique). Cette dernière est plus significative en termes de créations de connaissance "parce qu'elle se concentre sur la signification qui est transmise" (p.64). Les auteurs finalisent en affirmant que :

L'information est un flux de messages, au pas qui la connaissance est créée par ce même flux d'information, ancré dans les croyances et compromis de son détenteur. Cette compréhension souligne que la connaissance est essentiellement associée à l'action humaine (p.64).

Le défi de connaître l'environnement

Penteado (1996, p.32) a résumé les défis les plus pressants pour les entreprises et ses administrateurs, selon Reynolds (1992, pp.4-18).

1. Les systèmes supérieurs d'information différencient des entreprises en ce qui concerne le service au client ;
2. Des structures organisationnelles hiérarchiques sont défiées par une tendance à ce que les organisations deviennent chaque fois plus comme un réseau. Les organisations elles-mêmes s'unissent les unes aux autres comme des grands réseaux où des coalitions d'industries fonctionnent comme une équipe ;
3. Les entreprises ont besoin de devenir globales et d'organiser leurs travaux et leurs produits sur une base mondiale ;
4. Grâce aux nouvelles technologies, les entreprises sont capables de concourir dans les marchés globaux avec des produits personnalisés, adaptés aux préférences locales, sans perdre les bénéfices de l'économie d'échelle ;
5. Les entreprises doivent repenser leur manière de faire des affaires, en incluant tout, du flux de travail aux définitions de fonctions, procédures administratives, processus de contrôle et structures organisationnelles, en plus de sa culture et ses valeurs mêmes. La tendance est de donner du pouvoir à la personne qui sera responsable pour tout le processus, de son début jusqu'à la fin.
6. Les organisations doivent changer leurs croyances, conventions, cultures, structures et comportements pour capturer les bénéfices du processus de la ré ingénierie organisationnelle.

Penteado (1996) souligne que "la révolution dans la communication affecte aussi la manière dans laquelle les organisations dessinent leurs stratégies de négoce" (p.33). Il cite Hamel et Prahalad (1994a) qui soutiennent le développement d'un nouveau concept de stratégie, basé sur une connaissance profonde de son propre négoce et de son environnement. Dans cette nouvelle ère, ils affirment que, "n'importe quelle entreprise incapable d'imaginer son futur, ne sera pas vivante pour l'éprouver" (p.xii). Ceci est vrai parce que se charger de moderniser et améliorer son administration, n'est pas l'assurance de que cette organisation puisse restaurer son hégémonie sur une industrie ou qu'elle interceptera le futur. Elle serait seulement se chargeant de ses affaires courantes.

Ainsi, Hamel et Prahalad (1994b) insistent sur le fait que la tâche primordiale des administrateurs aujourd'hui est de construire une vision prospective de tout son industrie, ce qui est équivalent à une profonde connaissance des

tendances et discontinuités - technologiques, démographiques, réglementaires ou de la façon de vivre - utilisées pour transformer les frontières de l'industrie et créer un nouvel espace de compétition. Cette vision de l'industrie permet à une entreprise d'arriver au futur en première place en éliminant une position établie de prépondérance. Elle montre où l'entreprise s'oriente et permet qu'elle-même soit au contrôle de l'évolution de son industrie et par cette voie, de son propre destin (p.65).

Les auteurs critiquent aussi la "mode" d'être orienté vers les clients. Ils argumentent que les clients "sont notoirement faibles à pressentir le futur" (p.67). Ainsi, les organisations qui cherchent uniquement à servir aux nécessités articulées par ses clients "cèdent des vastes opportunités à des compétiteurs plus prévoyants" (p.67).

Pourtant, pour être victorieuses les entreprises auront besoin d'une vision stratégique qui privilégie l'habileté à construire des coalitions industrielles et à travailler en équipe. Ceci concerne aussi un compromis d'obtenir de nouvelles compétences, d'innover, de trouver la configuration correcte d'un service ou produit dans le marché, d'établir ses propres normes comme les normes de l'industrie et gagner la bataille permanente pour la participation au marché. Tout cela oblige à connaître son propre négoce et, au-dessus de tout, son environnement.

Penteado (1996, pp.34-35) indique que pour Russel et Russel (1992),

la pratique de 'corporate entrepreneuriat' peut être vue comme l'amélioration des compétences organisationnelles grâce à l'innovation" (p.639). Ils lient cette définition avec celle de Van de Ven et Ferry (1980) où l'innovation est le développement et l'implémentation de nouveaux concepts qui peuvent être des innovations techniques, des produits, des processus et/ou de l'administration. Aujourd'hui les activités d'entreprise sont agrandies pour inclure l'implémentation d'un "vaste champ d'innovations comme manière de créer et d'explorer des opportunités apparues dans les environnements compétitifs" (p.640). Les auteurs distinguent entre l'administration d'entreprises pour innover et l'administration stratégique pour innover. L'administration d'entreprises pour innover inclut "l'engagement d'entreprises et des efforts individuels mais inhabituels pour obtenir une supériorité par la voie des innovations bien comme des engagements dans des efforts persistants pour innover" (p.640). L'administration stratégique pour innover concerne une "persistante tendance sanctionnée par l'articulation d'activités ayant relation avec des innovations et l'allocation de ressources pour ce but, qui est une des composantes de la stratégie

*de l'entreprise" (p.640). Mitchell (1989) dit que les organisations qui formalisent ces processus innovateurs comme une stratégie d'entreprise, comme la 3M, la **Johnson & Johnson** ou la **General Electric**, sont capables de la maintenir même quand l'administrateur responsable du processus quitte l'entreprise.*

L'environnement, défini comme la somme de toutes les conditions et forces qui affectent les options stratégiques d'une organisation, mais dont le contrôle repose typiquement au-delà de son habileté est, comme décrit Ansoff (1979), une des principales priorités de l'entreprise moderne.

Penteado (1996, pp.36-37) souligne :

*Ansoff soutient que de 1800 jusqu'aux années 1950, les objectifs, tant des organisations qui recherchent des profits que des organisations qui ne visent pas à des profits (ONPs), étaient persistants clairs et stables. Chaque type d'organisation avait sa fonction de spécialisation et elles étaient séparées et vivaient dans des mondes différents. Les entreprises, Ansoff prend note, "étaient destinées à créer de la richesse nationale et à approvisionner des biens et services en tous les secteurs de la société où ils furent nécessaires et où il soit possible de générer des profits" (p.13). Les ONPs "se destinaient à la prestation de services sociaux qui étaient inhérentement non profitables" (p.13). En suivant les règles du jeu, les subsides du gouvernement ou des contributions de charité financiarisent les activités des ONPs. Du côté des entreprises la philosophie était fondée sur Adam Smith et le laissez-faire. "Liberté sans restrictions pour concourir et pour s'engager dans quelconque activité commerciale" (p.13). La clarté des règles du jeu a contribué à ce que les "activités politiques et d'entreprise" deviennent moins prioritaires. Ansoff fait voir qu'à partir de la deuxième moitié du Siècle XIX, les effets collatéraux non désirés de la doctrine du laissez-faire ont entraîné les sociétés à "renforcer des limites au comportement des entreprises" (p.13). Telle action a été combattue par un effort direct de marketing politique, avec les entreprises en unissant leurs forces en des associations pour défendre la cause de la libre entreprise. Néanmoins, "dans la seconde moitié du Siècle XX, la situation a commencé à changer radicalement. Pas seulement les règles du jeu, mais aussi la raison d'être des entreprises et ONPs commence à être contestée et a changé pour toutes les deux" (p.13). Ansoff conclut donc que les organisations s'engageront chaque fois plus dans des activités politiques et d'entreprise et dans le marketing politique et souligne-t-il, "aux États-Unis le marketing politique est appelé **lobbying** ou relations publiques" (p.12).*

Et Penteado continue (1996) :

Pour Higgins et Vincze (1989), la nouvelle tendance des affaires va dans le sens d'une réglementation croissante et d'un contrôle par les gouvernements nationaux, des régions, des provinces et des cantons. L'environnement de l'entreprise doit donc être constamment observée et les organisations doivent s'adapter ou essayer de le changer" (p.84). Les auteurs argumentent qui pour faire face à son environnement, les organisations doivent s'engager à des actions

politiques et de lobby légal et socialement responsables. Ils sont emphatiques à propos de la responsabilité sociale. L'organisation doit viser les nécessités de la société et la qualité de vie comme des valeurs stratégiques. "Celle-ci n'est pas seulement une position morale, c'est la dure réalité. Lorsque les entreprises ne donnent pas assez d'attention à ces nécessités, elles sont fréquemment obligées par la société et par le gouvernement à se comporter d'une manière qui à la fin des comptes devient plus coûteuse et moins profitable" (p.87). Il ne s'intéressent pas à la dimension morale. Ces forces externes sont des limites réelles pour les organisations. Elles doivent donc avoir des stratégies pour traiter avec ces forces. "Les entreprises doivent explorer de manière continue les thèmes et les sujets concernant la responsabilité sociale pour déterminer leurs impacts sur la société. Aucune surprise ne doit arriver" (p.94).

Penteado (1996, p.37) signale que l'argument principal de Higgins et Vincze (1989) est que les organisations ont besoin de faire des analyses de l'environnement externe qui prennent en compte des facteurs comme

le gouvernement, la technologie, la concurrence, l'industrie, des nouveaux concurrents, consommateurs et clients, les ressources naturelles, la scène internationale, les groupes de pression et les créanciers, d'entre autres. Le propos d'une telle analyse est "d'accéder à l'environnement et déterminer l'adéquation entre cet environnement, ses menaces et opportunités et les forces et faiblesses révélées par le profil organisationnel développé avec l'analyse stratégique (p.82).

Avec un pouvoir de synthèse propre à ses activités de normalisation, l'Association française de normalisation, AFNOR (1998) explique que

les organisations oeuvrent dans un environnement de plus en plus complexe caractérisé par une pression concurrentielle de plus en plus dure, une globalisation et une mondialisation des échanges, une forte évolution soit sur le plan technologique soit sur le plan réglementaire ou économique, une accélération des flux d'information, une obligation d'anticipation, d'innovation, de réactivité et de prise de décision rapide. Ce nouvel environnement est un signe de la naissance de la société de l'information dans laquelle gestion, qualité, rapidité de l'information sont des facteurs déterminants de compétitivité (p.4).

Intelligence et Veille

Le décret 4376/2002, du 13 septembre, 2002, du gouvernement du Brésil définit "intelligence" comme

l'activité d'obtention et d'analyse de données et informations, de production et diffusion de connaissances, à l'intérieur et à l'extérieur du territoire national, relativement à des faits et à des situations d'influence immédiate ou potentielle sur

le processus décisionnel, l'action gouvernementale, la sauvegarde et la sécurité de la société et de l'État (p.1).

Le même document considère "contre-intelligence"

l'activité dont l'objectif de prévenir, détecter, obstruer et neutraliser l'intelligence adverse et des actions de n'importe quelle nature que constituent une menace à la sauvegarde des données, des informations et des connaissances d'intérêt ou de la sécurité de la société et de l'État, aussi bien des champs et moyens qui les retiennent ou dans lesquels ils transitent (p.1).

À enregistrer qu'il existe des caractéristiques propres du mot "intelligence" selon la langue utilisée. En anglais ou en langues d'origine latine comme le portugais, par exemple. Le dictionnaire Aurélio (Ferreira, 1986) mentionne huit sens pour ce mot :

1. *Faculté d'apprendre, d'appréhender ou de comprendre ; Perception, compréhension, intellect, intellectualité ;*
2. *Qualité ou capacité de comprendre et de s'adapter facilement ; Capacité, pénétration, acuité, perspicacité ;*
3. *Façon de comprendre ou d'interpréter ; Interprétation ;*
4. *Accord, harmonie, entente réciproque ;*
5. *Relations ou accords secrets ; Connivence, machination, trame ;*
6. *Dextérité mentale ; Habileté ;*
7. *Psycho. Capacité de résoudre des nouvelles situations problématiques par l'intermédiaire d'une restructuration des données perceptives ;*
8. *Personne intelligente (p.774).*

Par contre, le dictionnaire Webster (1994) indique que "intelligence" est :

1. *Capacité de raisonner, de comprendre et des façons similaires d'activité mentale;*
2. *Manifestation de cette capacité ;*
3. *Faculté de comprendre ;*
4. *Connaissance d'un événement, circonstance, entre autres, reçue ou transmise ; Nouvelle, information ;*
5. *Réunion ou distribution d'information, ESP. Information secrète ;*
6. *Gouv. a) Information sur un ennemi ou un potentiel ennemi. b) La conclusion validée retirée de cette information. c) Une organisation ou agence employée pour réunir ces informations : Intelligence militaire ; Intelligence navale ;*
7. *(Fréq. cap.) Un être intelligent, ESP. Un être incorporel ;*
8. *Échange d'information (...) (p.739).*

La langue anglaise inclut le sixième sens (littéral, sans allusion métaphorique) du Webster (1994), d'information sur un ennemi réel ou potentiel. Herring (1996), par exemple, classifie intelligence comme "la connaissance de l'environnement compétitif de l'organisation et de son macro-environnement, appliquée à des processus de prise de décision, aux niveaux stratégique et tactique". Herring considère un système d'intelligence compétitive "le processus organisationnel de collecte et analyse systématique de l'information qui, de son tour, est disséminée comme intelligence à ces utilisateurs en appui à la prise de décision, ayant en vue la génération ou sustentation d'avantages compétitifs".

Le site du Canadian Institute for Scientific and Technical Information et de Federal Partners in Technology Transfer (2001) rapporte une définition bien succincte d'intelligence compétitive par Jerry P. Miller :

Le processus qui consiste, en suivant les principes éthiques, à collecter, analyser et disséminer de l'intelligence correcte, pertinente, spécifique, opportune, avec une vision prospective et actionnable, en concernant les implications de l'environnement des affaires, des concurrents et de l'organisation elle-même.

Une autre définition du site est celle d'intelligence compétitive technique (ICT), de Jonathan Calof :

L'ICT s'occupe des aspects techniques et scientifiques de l'environnement. Elle englobe le principe selon lequel l'avantage compétitif ira aux entreprises qui savent où les technologies du futur sont créées aujourd'hui.

La norme AFNOR XP X 50-053 de 1998 établit que l'intelligence compétitive économique est

La totalité des actions coordonnées de recherche, traitement et distribution en vue de son exploitation, de l'information utile aux acteurs économiques. Ces diverses actions sont menées légalement avec toutes les garanties de protection nécessaires à la préservation du patrimoine de l'organisme, dans les meilleures conditions de qualité, de délais et de coût (p.6).

La norme AFNOR XP X 50-053 a défini "veille" comme :

L'activité continue et en grande partie itérative visant à une surveillance active de l'environnement technologique, commercial, entre autres, pour anticiper les évolutions (p.6).

Selon Henri Martre (1994, apud. Masson 2001, p.25) "la notion d'intelligence économique" implique avoir "une intention stratégique et tactique" et "une interaction entre tous les niveaux de l'activité" responsables pour cette fonction :

Depuis la base (interne à l'entreprise), en passant par les niveaux intermédiaires (interprofessionnels locaux), jusqu'aux niveaux nationaux (stratégies concertées entre les différents centres de décision), transnationaux (groupes multinationaux) ou internationaux (stratégies d'influence des États-nations) (pp.16-17).

Masson (2001) rapporte que Dedijer considère la définition française (AFNOR, 1998) du terme "intelligence économique "

un modèle à suivre pour les autres pays [puisqu'elle] cherche à répondre à la problématique des 'relations entre le gouvernement, l'intelligence et la société en général et la privatisation de l'intelligence en particulier dont l'aspect le plus crucial est comment les entreprises et d'autres systèmes sociaux utilisent-ils les moyens d'intelligence gouvernementale et ses produits' (p.212).

Dedijer et Jequier (1987) entendent que cette évolution enregistrée en France décrit un nouveau type d'intelligence, "l'intelligence sociale", définie comme le "processus d'acquisition de l'information dans son sens le plus ample, par une société ou une organisation qui la **traite, évalue, stocke et l'utilise pour l'action**" (p.13).

Ce travail adopte la vision française en privilégiant la notion d'intelligence économique et les notions de veille et d'intelligence sociale selon les définitions en haut.

Dimension économique de l'intelligence

Dedijer (2003) et Steele (2005) mentionnent que les États-nations ont toujours investi beaucoup en intelligence de l'État sans, néanmoins, avoir conjuré le spectre des fautes constantes justement dans son appareil d'**analyse**. Néanmoins l'un des développements le plus important de la société de l'information est que chaque fois plus d'individus prennent conscience qu'une information qui répond à une question stratégique de manière synthétique, complète, valide, confidentielle et relativement récente, a beaucoup de valeur et, le plus intéressant encore, cette valeur augmente avec son utilisation. Fayard (2000) explique "qu'en différence avec des nombreux biens matériels, l'information (...) et la connaissance ne déperit pas quand elle est utilisée, bien au contraire" (p.128). Il continue,

consommer, échanger, traiter l'information est, à l'inverse, le gage de son enrichissement et de son intégration sous la forme de connaissance. Sa valeur s'accroît en fonction de l'utilisation et décroît quand isolée et considérée comme un stock sûr, solide et définitif. Le flux et la confrontation l'enrichissent pendant que l'immobilité l'appauvrit et ceci s'accroît dans un monde vélocé dans lequel l'obsolescence menace en permanence (p.128).

Cependant, développer des nouvelles compétences et des structures organisationnelles ou sociales pour qu'on puisse identifier, classer et disséminer l'information nécessaire exige une vision stratégique, du temps et des ressources. Par exemple : L'accès aux grandes bases de données internationales n'est pas bon marché. Marchand (1999) nous donne une idée du coût d'accès à la base de l'**Institute for Scientific Information (ISI)**. L'accès à jusqu'à 5000 références, coûtait un dollar par référence. Cette valeur diminue progressivement mais continue d'être relativement élevé pour une micro ou petite entreprise. Jusqu'à 50.000 références le coût est de US\$.30 cents par référence, 100.000, US\$.20 cents et à partir de 1.000.000 références, US\$.10 cents.

Dedijer (2003, p.5) indique que les dépenses dans le domaine de l'intelligence et de la sécurité dans le monde, en 2003, étaient de l'ordre de 150 milliards de dollars. Il rappelle une étude d'IBM qui évaluait les investissements en "**business intelligence & security** à 70 milliards de dollars, parmi lesquels, 30 milliards concernent seulement les États-Unis, 1,2 milliards le Royaume Uni, 5 milliards chacun pour la Chine et la Russie et les presque 20 milliards de dollars restants à tous les autres pays du monde" (p.5). Steele (2005, v-1, p.7), pour sa part, calcule que le budget d'intelligence des États-Unis serait dans les environs de 50 milliards de dollars.

Dedijer (1984) mentionne que le gouvernement suédois a conclu en 1976 que son "intelligence devrait être réorientée des préoccupations exclusivement militaires pour les problèmes du développement et la sécurité du pays" (p.33). En Suède, "en 1960, 65% de dépenses en intelligence et sécurité étaient dans la sécurité militaire et 15% étaient en l'intelligence technique et économique. En 1985, l'intelligence militaire représentait 30% du budget et l'intelligence économique et technique 60%" (Dedijer, 2003, p.19).

Downham et al. (2002, p.161) affirment que "les complexités et les pressions de la concurrence globale et la nécessité de l'intelligence compétitive sont croissantes". Ils expliquent que le service d'intelligence agrège de la valeur. En disséquant les concurrents et en découvrant leurs points faibles, il pointe aussi les faiblesses de l'organisation. Ils mentionnent d'autres retours importants : Prévoir les changements du marché, prévoir les attitudes de la concurrence, découvrir des concurrents nouveaux et/ou potentiels, apprendre avec le succès et l'échec des autres, augmenter l'extension et la qualité des acquisitions et/ou des partenariats, apprendre sur les nouvelles technologies, produits et processus qu'affectent

les affaires et apprendre sur les changements politiques, législatifs et normatifs qui puissent affecter le négoce.

Une consolidation de cinq études sur l'organisation de la fonction d'intelligence en presque 600 entreprises faite par Laackman et al. (2002, p.230) a indiqué que les services d'intelligence utilisent en moyenne 3,8 employés et leur budget peut varier entre 350 mille et 651 mille dollars, en fonction du genre d'activités et des objectifs de l'intelligence, entre autres raisons.

Bryant et al. (2002, p.351) calculent, pour leur part, que pour commencer une opération d'intelligence il faut compter avec de 400 mille à 450 mille dollars, pour avoir une équipe "formelle" avec "un gérant, deux spécialistes et du personnel d'appui", en ajoutant rémunération, systèmes d'information, accès à des informations (bases de données), signature de plusieurs services et assistance technique interne aux ordinateurs.

À la case ressources humaines, il faut inclure non seulement l'entraînement mais aussi les salaires et le temps dédié à l'activité. Jakobiak (1991, p.184) indique que dans une organisation de jusqu'à 500 personnes, le système d'intelligence peut être l'affaire d'un seul technicien appartenant à la Direction Générale, avec dévouement exclusif et une secrétaire technique. Le responsable peut créer un réseau d'observation et constituer et animer un réseau de spécialistes et analystes avec trois ou quatre personnes spécialisées dans les sujets de veille. Il peut aussi élaborer, gérer et maintenir actuels les rapports d'intelligence, en assurant sa mémorisation et son archive. Pour une petite entreprise de 50 personnes Jakobiak recommande de sous-traiter le travail d'intelligence à une tierce partie, avec l'administration du contrat par un directeur ou par son adjoint.

Quand il n'est pas possible d'avoir une équipe formelle, Bryant et al. (2002) recommandent d'agir en premier lieu pour démontrer l'efficacité des services d'intelligence. Faire une analyse des nécessités ; "Découvrir ce qui est en train de retirer le sommeil des dirigeants" (p.354) et travailler après pour générer des solutions pour ces questions cruciales pour l'organisation. L'équipe doit chercher à mesurer les bénéfices de l'activité et à les traduire en ressources. Bryant et al. (2002, p.355) indiquent "quatre mesures quantitatives basiques" pour l'activité d'intelligence : Economiser du temps, réduire des coûts, éviter des dépenses et améliorer les recettes.

Rosenkrans Junior (2002) rapporte que la question principale quant à la valeur des activités d'intelligence est "sa contribution à la stratégie et à la décision qui est prise". Pour mesurer la valeur de l'intelligence générée il recommande d'utiliser trois types d'indicateurs : Quantitatifs, d'innovation et de relation avec la stratégie. Les indicateurs quantitatifs démontrent "ce qui à été affecté par l'activité, l'opportunité, la valeur et le coût". Les indicateurs d'innovation mesurent comment l'intelligence "supporte le processus d'innovation dans l'organisation" et les indicateurs de relation avec la stratégie révèlent la contribution de l'intelligence "non seulement pour résoudre le problème présent, mais pour soutenir la stratégie" (p.342).

Bryant et al. (2002) expliquent que certaines activités d'intelligence comme le montage d'un système d'information pour la collecte, traitement et stockage des données "doivent être financées par la branche centrale" (p.356) quand elles bénéficient à toute l'organisation. Ils alertent que "quand l'organisation est trop décentralisée" (p.356) il est important de maintenir une fonction centrale de coordination du service d'intelligence, soit en contribuant pour le financement des infrastructures et pour accomplir les fonctions de veille de l'environnement extérieur et des scénarios futurs, soit pour agir comme lieu de rencontre et de facilitation de la dissémination des outils et techniques de traitement et d'analyse de l'information.

Galvin (2002) observe qu'à Motorola il y a un service central avec 10 personnes intégrant la "stratégie corporative" (p.146). Le service prend en charge toute l'organisation, en collectant et en traitant des informations. Néanmoins, chaque division et département désignent "une, deux ou trois personnes" (p.146) pour faire une interface avec le service d'intelligence et être aussi les multiplicateurs de ces activités dans les divisions et départements.

Rouach (1999) indique qu'au Japon, environ 1,7% du budget des entreprises est investi en intelligence. Les employés dédient 5% de leur temps au traitement et à la diffusion de l'information collectée dans l'environnement interne. Rouach (1999, p.52) cite Villain (1991) qui évalue les investissements annuels du gouvernement japonais en information entre 4 et 6 milliards de dollars. L'information scientifique et technique reçoit de 10 à 15% du budget totale de recherche et développement.

Quelques chiffres sur l'effort des entreprises japonaises ont été recueillis par Rouach (1999). Chez Mitsui il y a 162 bureaux dans le monde avec un réseau de télécommunication similaire à des petits pays tel que la Belgique, Suisse ou Hollande. Sa devise est "l'information est le

sang de l'entreprise" (p.56) et 84 personnes font partie du département qui s'occupe de la veille et du transfert des technologies étrangères. Chez Nippon Electric Company (NEC) il y a 250 personnes que s'occupent de l'information. "En 1982, 150 employés de la NEC eurent pour seule mission de voyager par-tout le monde pour détecter des innovations intéressantes" (p.57). Chez NEC, 300 personnes s'en chargent de l'intelligence technologique. Chez Mitsubishi, 30 personnes s'occupent uniquement de la veille des brevets des concurrents.

Rouach (1999, p.94) rapporte qu'en France, le groupe Hutchinson, leader en la transformation du caoutchouc industriel en Europe, applique 7% du budget de recherche et développement à l'intelligence en trois thèmes : Matériels, Processus et Produits. Le Groupe CFPI, spécialisé en produits sanitaires et dans la chimie fine (Rouach, 1999, p.95), destine 5% du budget de recherche et développement à l'intelligence concurrentielle et commerciale. Entre les résultats obtenus, le développement de produits originaux et aussi des accords de licences :

L'entreprise cite l'exemple d'utilisation d'un rapport des responsables de la veille concernant un désinfectant inédit développé par une société américaine et pouvant être appliqué pour l'hygiène des mamelles des vaches laitières. Des contacts ont été pris et [il a été calculé qu'] à terme cette innovation pourrait se traduire par un chiffre d'affaires de près de 10% des ventes en agroalimentaire. Un accord de licence a été signé avec le laboratoire américain.

Selon Rouach (1999), "la valeur ajoutée de la veille consiste à donner des informations valides à la direction des entreprises pour aider à la prise des décisions stratégiques" (p.120). L'auteur raconte "l'expérience de la SNECMA" (p.35), fabricant français de moteurs d'avions à réaction. Un nouveau matériel a été créé aux États-Unis, le carbone-carbone. La SNECMA découvrit,

s'intéresse, se renseigne, visite, manifeste son désir d'acheter, croise les infos techniques, imagine des applications, croise les infos économiques. Le tout rapidement (un peu superficiellement). Avant que cette technologie ne soit prise. Avant qu'elle ne soit trop chère. Avant qu'elle ne soit considérée stratégique et interdite à la vente, ce qui devait arriver deux ans après (Rouach, 1999, pp.35-36).

Rouach narre que la SNECMA a acheté la technologie et sa foi était encore plus forte que celle des découvreurs d'outre Atlantique. La nouvelle technologie a été combinée avec les recherches internes. Des démonstrations ont été faites de divers nouveaux produits, de tuyères de propulseurs à disques de freins d'avions. Comme le marché américain était plus large que le français, la SNECMA mit tout son influence de marque pour patronner la nouvelle

technologie aux États-Unis ; Elle a organisé des partenariats commerciaux et technologiques, vendu licences "dix fois plus chères que celles acquises dix ans avant, mais cette fois ciblées, limitées, contrôlées" (Rouach, 1999, p.36). L'entreprise française "devient leader mondial, fournisseur apprécié, partenaire désiré". Créa "un nouveau métier : La carbonurgie" (p.36) avec une croissance annuelle de 25%. À la fin du siècle, la SNECMA contrôlait le tiers du marché mondial, produisait 300 tonnes par an de la fibre à carbone en "trois usines, dont une aux Etats-Unis" (p.36). Rouach (1999, p.36) conclut en citant Pierre Bertin, de la SNECMA :

un excellent retour sur investissement et une magnifique affirmation française..[Un résultat d'] un croisement constant de l'innovation scientifique, technique et industrielle, commerciale et de l'intelligence économique, des matériaux, produits, marchés et acteurs (p.36).

Dimension sociale de l'intelligence

La vision selon laquelle l'intelligence apporte des avantages compétitifs et économiques pour les organisations et États n'aurait pas beaucoup de signification si elle restait limitée à cette approche mercantiliste. Un courant de pensée dont l'origine est attribué par le professeur Stevan Dedijer, de l'Université de Lund, Suède (Dedijer, 1984), au philosophe pragmatique américain, John Dewey (1910, 1927, 1938, 1939) est en train d'acquérir beaucoup de popularité entre les auteurs et spécialistes du terrain. Il affirme que l'humanité est "un système social interconnecté" (p.20) faisant face à des problèmes globaux tels que la croissance de la population mondiale, la limitation des ressources de l'eau potable, la pollution, l'échauffement globale et les armes nucléaires, biologiques et d'autres. Et, continua Dedijer (1984) :

Dans des crises comme celles mentionnées, ces systèmes interdépendants et interactifs, tout être humain individuellement, chaque organisation, gouvernement et État recherche sa propre immunité contre la destruction, le maintien de ses qualités de subsistance, de sa sécurité et à rester libre de l'appréhension et du danger. (...) En ces crises, l'homme regarde à soi même et aux "mondes" à son tour avec un désir d'identifier des données, des faits, des questions, des informations et de la connaissance de tendances qui puissent offrir des opportunités pour sa survie, sa sécurité, sa stabilité et son développement (p.22).

Dedijer explique que la démocratie "devient une force productive, une nécessité basique" en fonction de la "démocratie de l'information" (Dedijer, 1979a, p.529), le droit des citoyens d'un pays "d'obtenir, produire, distribuer et garder information". Ce droit est, aujourd'hui essentiel

pour qu'un pays "obtienne et créé des connaissances scientifiques, d'autres domaines et aussi du savoir-faire" (p.529).

Dedijer (1979b) argumente aussi que, dans les pays en développement, nous devons faire attention au développement sociale, politique, économique, culturel et à d'autres contextes, en relation à la science et à la technologie, comme à la recherche

d'innovations sociales spécifiques telles que des nouvelles lois, des organisations, des professions, des codes de conduite, des normes de comportement, des systèmes intelligents, des normes pour des subsides et d'autres initiatives (...) pour importer, adapter ou créer des technologies étrangères ou nationales de manière à contribuer le maximum possible pour ses nécessités spécifiques et ses objectifs de développement (1195).

Cette vision est partagée par le professeur Henri Dou (2001b) qu'a été un des responsables par le premier cours de spécialisation en Intelligence Compétitive au Brésil. Pour lui

l'intelligence organisationnelle et compétitive est valide seulement si les concepts, les méthodes et les opportunités en jeu sont clairement compris par ces qui agissent dans le développement du pays. En même temps, nous ne pourrons changer de paradigme et dominer ces changements que si nous nous approprions nous-mêmes des méthodes et établissons rapidement une intelligence et une organisation 'à la Brésilienne', en prenant en compte le passé, les mentalités, les moyens et la culture de ce pays" (pp.27-28).

Comme "on n'arriverait pas à l'intelligence par l'accès passif à l'information ou même, seulement par les connaissances transmises par les écoles et universités", (p.28) elle **doit être créée**, indique Dou (2001b) "c'est au long de ce processus de création qu'un système utile au pays va en s'élaborant, intégré dans sa culture et dans ces scénarios et tournés vers le futur" (p.28). Dou (2001b) souligne que l'étape décisive pour la création "d'une intelligence brésilienne" (p.29) est celle de la capacité, que nécessitera de plus de livres, "d'instruments et des réflexions, aussi bien que de collaborations d'origine nationale et internationale" (p.30).

Un processus qui, selon Dou, agisse sur les méthodes de travail, de production, de recherche et d'utilisation des informations, en valorisant la synergie au lieu de la compétition et en mettant en évidence la nécessité de reconfigurer les structures organisationnelles. Le professeur de l'Université d'Aix Marseille III indique encore que comme la production d'information à un cycle chaque fois plus rapide, le pays nécessitera de "nouvelles technologies, nouvelles contributions de la pensée et nouveaux modèles mentaux" (p.29).

Dedijer (2003) signale que l'accélération constante des changements et des capacités d'échanger l'information est un processus mondial qui concerne "toute l'humanité" (p.2) qui est propulsé par l'Internet, laquelle infrastructure fait arriver ce que Robert Wright décrit en "**Nonzero : The logic of Human Destiny**" (Wright, 2000) :

*Avec le progrès de l'histoire, les êtres-humains iront chaque fois plus jouer des jeux à somme non-nulle avec de plus en plus des êtres humains.
L'interdépendance s'accroît et la complexité sociale s'accroît en portée et profondeur (Wright, 2000, apud. Dedijer, 2003, p.3).*

Masson (2001) note que Blaise Cronin et Elisabeth Davenport (1993) ont cité Stevan Dedijer comme un des précurseurs de cet

abordage intégré et holistique de l'intelligence. Il parle avec confiance d'une révolution de l'intelligence et de l'émergence d'une science de l'intelligence qu'intégreraient les chercheurs sur tous les aspects de l'intelligence : Biologique, individuelle, artificielle, gouvernementale (pp.139-140).

Pour Dedijer, l'information, la technologie et la connaissance sont la source réelle de la puissance des pays développés, au même niveau des ressources naturelles, le territoire ou l'économie et contribuent, selon Baumard (1991, p.46) à former une vision globale pour la compréhension des problèmes des hommes, des entreprises, des sociétés et des nations. Dans la mesure où de plus en plus d'organisations humaines pratiquent et mobilisent toutes les "intelligences", la "révolution" se passera inévitablement et contribuera pour résoudre pacifiquement les problèmes du monde.

Dedijer et Jequier (1987, apud. Masson, 2001, pp.142-143) expliquent que

les agences gouvernementales, les industries, les partis politiques, les syndicats, les armées, les groupes de pression politique sont tous impliqués en travaux d'intelligence. Ces activités, qui consistent en identifier des problèmes, des menaces, des opportunités et des défis, à réunir des vastes quantités d'informations et à les utiliser pour atteindre des objectifs fixés par les organisations elles-mêmes, sont très rarement reconnues comme des travaux d'intelligence. (...) L'intelligence est une activité importante de toute société, quel que soit son niveau de développement. En ce qui concerne la société, ce terme est parfois équivalent à intelligence (au sens psychologique du terme) d'un être humain (p.23).

Dedijer (2003) rapporte que "quand il a commencé à étudier l'histoire de l'intelligence" il a découvert qu'elle a eu "un rôle important dans l'histoire de la Suède". Entre autres points, il relève "une contribution unique" du pays baltique : Le **Saltsjöbad spirit**, créé en 1938, "une

innovation sociale pour assurer stabilité industrielle et éviter des conflits entre capital et travail" (p.12). Tous les ans, des représentants des patrons et des ouvriers se réunissent et cherchent un consensus sur "conditions de travail et salaires". Cette recherche concerne des discussions techniques sur des dossiers et des efforts pour "considérer les intérêts des deux cotés" (p.12).

Dedijer (2003, p.4) explique que le monde s'aperçoit "d'une révolution de l'intelligence". Il mentionne que le ex-chef de la **Central Intelligence Agency** (CIA), William Colby a été le premier à s'apercevoir, dans les années 70, de cette révolution. "Ses idées ont été suivies et développés par Wilhelm Arell et Stevan Dedijer" (p.4). Colby a énoncé cinq dimensions de cette "révolution de l'intelligence" :

1. Après la fin de la Deuxième Guerre Mondiale la majorité des Etats membres de l'ONU ont créé des "communautés de sécurité nationale et d'intelligence".
2. Les NTC et les nouvelles sciences comme la biotechnologie sont utilisées massivement "pour des recherches et des analyses pour de l'intelligence et de la sécurité".
3. Les entreprises et les autres organisations sociales utilisent aussi massivement l'intelligence pour "compétition, croissance et profit".
4. Des entités de représentation et contrôle sont créés avec la mission "d'assurer que la communauté d'intelligence et sécurité adhère aux principes basiques qui gouvernent la démocratie".
5. L'intelligence a un impact global. Elle devient "une ressource internationale", un instrument "pour aider l'humanité à identifier et résoudre ses problèmes par la négociation et la coopération au lieu de continuer à souffrir ou de lutter pour eux (W. Colby, Honorable Men : My Life in the CIA, 1978)". (Dedijer, 2003, p.4).

Dedijer ajoute trois autres dimensions à la "révolution de l'intelligence" :

6. Cette révolution est valide aussi en ce qui concerne les individus. Ils acquièrent "la capacité d'être inquisiteurs et d'obtenir des informations pour les aider à comprendre et à travailler leurs problèmes".
7. L'espionnage telle qu'il a été popularisé par les médias et le cinéma, comme une technique de guerre-froide, est "chaque fois moins appropriée pour les problèmes de l'intelligence actuels". Ses techniques deviennent "inadéquates et dépassées" parce que selon George Kennan, "plus de 95%" des informations nécessaires peuvent être obtenues "par l'étude soignée et compétente de sources d'information parfaitement légitimes" (p.4-5).
8. Une nouvelle "science du développement" a surgi avec une dimension holistique d'intégration des "connaissances d'individus de plusieurs systèmes sociaux en incluant des problèmes globaux et sciences naturelles" (Dedijer, 2003, pp.4-5).

Dedijer (2003) considère que "le fondateur de la philosophie pragmatique, John Dewey, a prédit que le développement de l'intelligence sociale allait éviter le conflit entre l'économie de

marché du laisser-faire et le socialisme totalitaire" (p.20). Et confirme : "**L'intelligence d'un système social dépend de l'analyse de l'information réunie**" (p.10).

Intelligence et communication

Grunig (1992a) introduit le concept de la théorie des jeux dans la pratique de la communication organisationnelle. Quand les programmes de communication adhèrent à des présuppositions de main double, qu'il a nommé "symétriques", ils deviennent des jeux à somme non-nulles : Tous peuvent gagner si le jeu est bien joué. Quand des programmes de communication pointent pour des présuppositions de main unique, nommés "asymétriques", ils deviennent des jeux à somme nulle. Il y a toujours un vainqueur et un vaincu.

Grunig et White (1992) remarquent qu'une mentalité de main unique domine la pratique et aussi les perceptions du public sur les relations publiques. Elle définit relations publiques comme "l'utilisation de la communication pour manipuler des publics en bénéfice de l'organisation" (p.40). Les auteurs proposent, en contre-proposition, que les programmes et campagnes de communication doivent chercher à rendre les organisations plus efficaces envers l'accomplissement de leurs missions et plus responsables envers la société, à mieux résoudre les problèmes de la société et les questions nationales et internationales.

Le principal propos des relations publiques de main double, selon Grunig et Repper (1992), est de "pourvoir une forme par laquelle des organisations et des groupes d'intérêt, dans un système pluraliste, interagissent pour administrer des conflits en bénéfice de tous" (p.9). Les relations publiques de main double servent l'organisation quand cherchent à suivre l'intérêt public, à développer une compréhension mutuelle entre les organisations et leurs publics et à contribuer à ce que les débats soient plus informés sur les questions de la société.

Pour Grunig (1992b) "le contenu symétrique et des communications ouvertes avec la haute administration marquent un renversement complet de la vision entrepreneuriale de l'époque du début des recherches de communication organisationnelle" (p.559). Il rend évident les présuppositions basiques de cette vision symétrique : La communication mène à l'entente, holisme, interdépendance, systèmes ouverts, équilibre constant, égalité, autonomie, innovation, décentralisation, responsabilité et résolution de conflits. Les théories pertinentes pour la communication de main double sont celles de la théorie des systèmes, du libéralisme, des groupes d'intérêt et de la négociation pour la résolution de conflits.

La vision de main unique mène les professionnels qui militent dans l'administration de la communication à agir d'une forme éthique, socialement irresponsable et même inefficace. Malgré leurs bonnes intentions il est difficile, sinon impossible, de pratiquer la communication d'une manière éthique et socialement responsable si on utilise des présuppositions de main unique ou asymétriques (Grunig et White, 1992, p.40).

La révolution de l'intelligence nécessite d'un certain type de communication capable de bâtir des consensus, des partenariats et du partage des informations et des connaissances pour résoudre des problèmes communs et ainsi permettre l'évolution de la société. Les caractéristiques de ce nouveau type de communication sont toutes des valeurs de la communication de main double, symétrique.

LA NÉCESSITÉ D'INFORMATION DES ORGANISATIONS

La nécessité d'information des organisations

Penteado (2005, pp.33-34) indique qu'

il y a une tendance universelle pour une "surenchère d'information". Naisbitt (apud. Faria et Quoniam, 2002) a été celui qui a résumé le mieux cette situation : 'Nous sommes noyés par l'information mais assoiffés de connaissance'. Faria et Quoniam (2002) signalent qu'à chaque année la connaissance humaine duplique en volume. Par an, sont édités dans le monde plus de 700 mille livres - presque 2 mille par jour - et de 100 à 300 mille revues scientifiques en apportant près de 10 millions de nouveaux articles scientifiques. Sont générées plus d'un million de brevets par an. À l'Internet existaient plus de 2,5 milliards de pages avec 550 milliards de documents, plus d'une dizaine de milliers de bases de données disponibles pour l'accès et nous comptons en centaines les différents outils de recherche.

Cette réalité a créé un dilemme et une nécessité qui contribuent pour l'inévitable développement de l'intelligence dans les organisations.

Le dilemme fait rapport aux différences entre l'information disponible et l'information nécessaire, ici nommée intelligence : 'l'information disponible est abondante, incomplète, douteuse, publique et confuse. Déjà l'information nécessaire est synthétique, complète, fiable, confidentielle et précise' (Faria et Quoniam, 2002). L'utilisation de l'information disponible est limitée. Plus elle est brute, plus le Décideur mettra du temps à se l'approprier. Pendant que pour tirer du sens de l'information brute il peut se passer plus d'un jour, l'information stratégique traitée et validée peut être entendue et appropriée par le Décideur en une minute. Quoniam (2001) signale qu'elle doit être "l'information correcte, à l'heure correcte, délivrée dans la forme correcte, à la personne correcte et doit résulter dans la décision correcte". L'utilisation est ce qui détermine la réelle valeur de l'information nécessaire. Si elle n'est pas utilisée, sa valeur est nulle.

La nécessité fait référence à ce que pour rester viables et saines, les organisations doivent dominer des méthodologies et techniques et appliquer des ressources humaines, financières et matérielles capables de réaliser les opérations exigées pour créer l'intelligence, soit, l'information nécessaire.

Pour cela, Levet (2001) indique que le problème n'est plus celui de l'accès à l'information qui est le ressort de la croissance et de l'emploi mais l'aptitude des acteurs à transformer, comprendre, interpréter et utiliser l'information' (p.38). Ceci a été reconnu par la communauté internationale dans des normes de qualité [...tel que] l'ISO 14.000, de 1996.

Dou (1999) observe que les organisations veulent s'anticiper en relation à ses concurrents. Elles nécessitent, pourtant, d'une information du présent ou du futur prochain et doivent l'utiliser le plus rapidement possible. Ce qui est en jeu est sa capacité d'analyser, en temps réel, le maximum d'informations possible et d'établir, à partir de cette masse, un nombre significatif d'intelligences économiques, financières, juridiques, diplomatiques, culturelles, sociales, scientifiques et politiques. Percevoir, avant ses concurrents, des signaux faibles, en indiquant une opportunité de négoce ou une technologie et agir dans ce sens peut déboucher sur de grands avantages compétitifs et maintenir la propre survie de l'organisation (Penteado, 2005, pp.33-34).

Dou (2001a, p.3) explique que l'inaction stratégique signifie rester en retard, ce qui peut devenir définitif. Le déphasage peut augmenter d'une manière très rapide, en provoquant une incapacité à concourir sur un marché déterminé, de la même forme que les avantages compétitifs s'y réduisent avec le temps. Dou (1999) indique que l'avantage concurrentiel consiste justement à rencontrer ce type de signal faible avant les autres et de l'utiliser pour orienter l'action, prévoir un mouvement stratégique d'un concurrent ou pour occuper un espace non occupé auparavant.

Boutin (2001) cite Morin (2000) en soulignant la nécessité que les organisations détectent l'émergent, les signaux faibles, nouveaux, sporadiques, isolés, qui sont peu significatifs. Après une période de temps, il apparaît une convergence significative des signaux, un début de quantification et le surgissement de tendances. Les nouvelles technologies se manifestent. Avec le passage du temps, ces technologies gagnent force et présence, deviennent des signaux organisés. À ce point nous pouvons affirmer aussi que si l'organisation ne prend pas connaissance de cette évolution, une menace est en train de s'installer. Quand cette technologie émergente dévient un signal forte, si l'organisation est "en dehors de la vague" (p.11), elle sera obligée à s'accommoder d'un fait concurrentiel. "A partir d'un certain moment" (p.11), conclut Boutin, "l'organisation perd son autonomie face aux événements mais elle garde, pendant ce temps, la liberté de s'associer ou d'abandonner un marché ou un produit" (p.11). Pour cela, continue Dou (2001b) "la gestion de l'information et son utilisation pour produire de l'intelligence deviennent des éléments basiques en favorisant le développement stratégique des organisations" (p.29).

En ce chapitre, nous traitons en détail de la morphologie et des méthodologies et techniques pour collecter, stocker, construire, associer, comprendre et utiliser des informations pour des fins opérationnels, de recherche, pour ajouter de la valeur et pour améliorer le processus de décision dans les organisations (Dou, 1999).

Morphologie de l'information

On considère que l'information brute doit obligatoirement être traitée pour devenir une information élaborée. Selon Boutin (2001, p.2), ces instruments pour traiter l'information brute sont des filtres du grand flux d'information qui permettent d'extraire trois types d'information : **L'information utile**, qui s'additionne aux processus de production de l'entreprise, **l'information stratégique**, qu'apporte un avantage compétitif à l'entreprise et **l'information fatale**, sans laquelle l'entreprise court le risque de disparaître.

Penteado (2001, pp.26-30) rapporte la typologie de l'information utilisée par Dou (1999), appuyée sur le modèle conçu par Hunt et Zartarian (1990). Les informations dans les organisations sont divisées en quatre grands groupes :

***L'information type texte ou formelle** - représente, en générale, 40% des informations disponibles dans les organisations. Elles sont des informations structurées, trouvées dans les bases de données, des livres, journaux, revues, CÉDÉROM, brevets et dans l'Internet.*

***L'information informelle ou fluide** - est une information bien moins structurée que la première. Elle représente presque 40% des informations disponibles dans une institution. En la majorité, elles proviennent de l'extérieur de l'organisation et obtenues par moyen de contacts avec des fournisseurs, des personnes invitées pour des congrès et des conférences. Pour les capturer, les organisations doivent appliquer des stratégies de réseau et rechercher l'intégration de toutes les sources disponibles de l'information.*

***L'information des spécialistes** - c'est la mémoire de l'entreprise, la connaissance de ces spécialistes, leurs situations et localisations. Elle représente à peu près 10% du total des informations qui circulent dans l'entreprise.*

***L'information des foires et congrès** - représente 10% de l'information disponible. Provient de la participation de personnes de l'organisation en des foires et congrès. La recherche des informations, les visites aux stands, la collecte de prospectus, peuvent montrer les tendances des concurrents dans des domaines d'intérêt spécifique. D'autres façons de 'découvrir' les secrets des organisations sont les articles présentés dans les congrès, les annonces d'emplois, les discussions dans les avions, les fournisseurs, les conversations avec des faux-amis et les **press-releases** de l'entreprise visée.*

Coelho, 1999 (apud. Penteado, 2001, p.28) présente deux tableaux avec les sources formelles et informelles de l'information.

Tableau 1 : Sources formelles d'information (Coelho, 1999, apud. Penteado, 2001, p.28)

Sources Formelles	
Patents	Rapports
Livres	Catalogues
Normes	Films
Legislation	Journaux
Revue	Bases de données
Actes de Congrès	Revue de la presse
Directoires	Internet

Tableau 2 : Sources informelles d'information (Coelho, 1999, apud. Penteado, 2001, p.28)

Sources Informelles	
Connaissance de l'entreprise	Prestataires de services
Fournisseurs	Étudiants et stagiaires
Expositions et foires	Candidats à un emploi
Missions et voyages d'étude	Congrès, Séminaires, Clubs
Comités	Reseaux personnels

Coelho (1999, apud. McGonagle Junior et Vella, 1996) explique que les sources informelles "tendent à donner des informations en relation avec le présent et le futur, pendant que les sources formelles reflètent plus le passé" (p.15). Des informations informelles "donnent plus d'opportunités pour localiser des informations ou des sources additionnelles" (p.15) pendant que les informations formelles sont plus "soumises aux limitations du temps (impression et distribution)" (p.15) finalement "les personnes ne sont pas filtrées au pas qui les sources formelles normalement le sont" (p.15).

La relation de l'information avec le temps

Une autre dimension à noter est celle du temps qui chaque personnage du système d'intelligence prend pour s'approprier de l'information. Selon Dou (1999) il est nécessaire de bien comprendre "la relation des informations avec le temps" (c2-p1) puis-qu'une bonne partie des problèmes qui se passent aujourd'hui dans les organisations pourraient être évités s'il y

avait conscience de cette liaison. Le temps a une relation très forte avec des informations qui se destinent à l'action. Par exemple, pour anticiper un mouvement d'un concurrent, plus l'information arrive vite au dirigeant responsable par l'action, plus grande sera l'avantage concurrentiel obtenu par l'organisation. Déjà des informations destinées à faire partie des collections d'une bibliothèque ou pour servir de base au curriculum de l'enseignement fondamental ont une relation relativement faible ou très ténue avec le temps.

Pour Dou (1999), la relation de l'information avec le temps a trois dimensions : Du formel à l'informel ; Du passé au présent et au futur ; Et de l'opérationnel à la stratégie. Quand nous analysons la première dimension, du formel à l'informel, nous pouvons considérer que l'information est 100% informelle quand nous sommes au niveau des personnes et 100% formelle quand nous sommes dans le domaine de la science. Quand nous passons du scientifique pour la technologie et, en suivant, pour le savoir-faire et le marché, l'information formelle diminue et l'information informelle augmente.

La deuxième dimension, du passé au présent et au futur, doit être examinée sous le point de vue de l'utilité. Plus l'information est formelle plus elle sera ancienne et représentative du passé. Les informations sur le présent et le futur sont informelles. L'entreprise, souligne Dou (1999) à besoin de s'anticiper en relation à ces concurrents. Elle a besoin pourtant d'une information du présent ou du futur prochain. Ainsi, une attention spéciale doit être conférée à la recherche des informations informelles. Ceci, explique Dou (1999), parce que dans la meilleure des hypothèses, quand nous recherchons une information dans une base de données scientifique, formelle celle-ci sera une information vieille d'au moins deux ans et deux mois.

Ce période correspond au temps minime qu'une information parcourt le "chemin de la Science" : La lecture par le chercheur pour appréhender une information, le premier pas, pourrait prendre un mois. Depuis un autre mois pour que le chercheur ait des idées, amenés ensuite à son laboratoire où il travaillera encore plus, au minimum, un an. En obtenant des résultats, le chercheur écrit et depuis publie, période qui prend plus un an. En additionnant tous ces temps, finalise Dou (1999), le temps minime passé entre la lecture qui a généré l'idée et sa publication dans une revue scientifique sera de deux ans et deux mois.

Ainsi, quand nous lisons un résultat de recherche ou une publication scientifique nous apprenons comment l'auteur pensait-il y a deux ou trois ans. Non comment il pense

maintenant. "L'information stratégique importante", continua Dou (1999), "est comment il pense maintenant".

La troisième dimension, de l'opérationnel à la stratégie, a à voir avec l'action. Dou (1999) rappelle qu'une information opérationnelle a un objectif défini, vise à réaliser un projet, un produit, une expérience. Elle se caractérise par son bas coût, quand entraîne des rapports réguliers et un système d'informations cohérentes. Les systèmes d'alerte qu'intègrent et interprètent des informations sur des thèmes spécifiques génèrent des informations intermédiaires entre la tactique et le stratégique. L'information utilisée dans des systèmes de planification stratégique est plus complète. Elles sont stratégiques puis-qu'elles concernent les menaces et opportunités, projettent des positions futures et indiquent des nouveaux concurrents et recommandent des réponses possibles.

Le cycle de l'intelligence

Faria et Quoniam (2002, p.10) décrivent graphiquement le processus de traitement automatique de l'information pour générer de l'intelligence. Dans ce cas, ils se réfèrent au cycle de l'intelligence. Voyez la Figure 1.

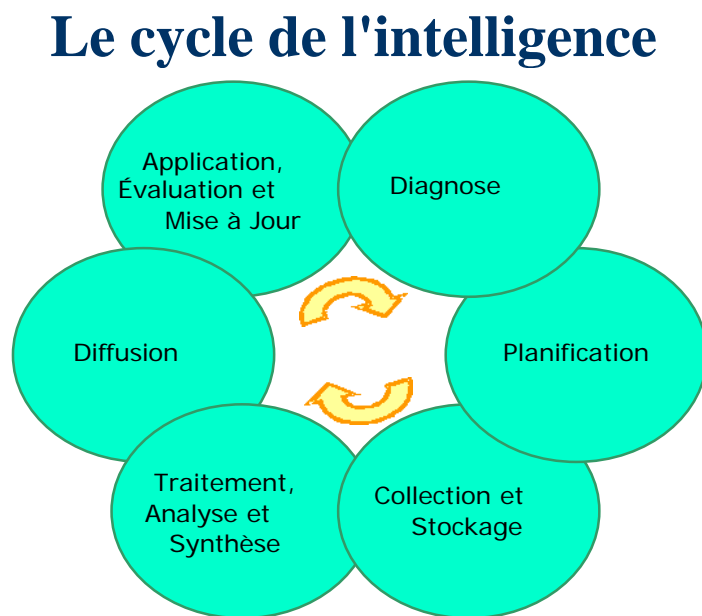


Figure 1 : Le processus de génération de l'intelligence (Faria et Quoniam, 2002, apud. AFNOR, 1998)

Le cycle a six étapes : Diagnostic ; Planification ; Collection et Stockage ; Traitement, Analyse et Synthèse ; Diffusion ; Application, Évaluation et Mise à Jour. C'est un processus général, utilisé à chaque opération. Il est le moteur principal des systèmes d'intelligence organisationnelle.

Sa productivité peut être accrue avec les nouvelles technologies de stockage de données, apparues récemment que ne dispensent pas la préparation et la reformatage des données et textes avant des analyses (Quoniam, 1996), réalisées en deux versantes ; L'analyse de données et l'analyse de textes.

Collection et stockage : Les entrepôts de données

Les bases de données sont formées par un ensemble de références et chaque registre est intégré par un ou plusieurs champs rangés selon une séquence fixe. Pour Penteadó (2005, p.36)

Quoniam (1999) avertit que, en la conquête de l'information nécessaire, une des premières providences serait l'intégration de toutes les informations en un système unique pour "information à l'intelligence". Cette nécessité d'intégrer les informations est un des facteurs qui a contribué au développement d'un nouveau concept d'organisation des données, les entrepôts de données (Datawarehouse-DW). Son origine a été les grandes chaînes de magasins des États-Unis (EUA). Pour rester un degré au-dessus de la concurrence, elles ont monté des entrepôts de données, en ajoutant des bases de données financières, juridiques, politiques, administratives, de ressources humaines, commerciales, techniques et scientifiques, d'entre autres. En termes techniques, l'entrepôt de données représente un espace informatique où les informations réunies des systèmes de données opérationnels deviennent disponibles pour accès en ligne. Deux environnements sont créés, en séparant, dans des différents ordinateurs, les systèmes en ligne opérationnels et les systèmes en ligne analytiques (p.36).

Jambu (2002, p.5) mentionne 12 règles pour avoir un bon DW : 1) Concept multidimensionnel ; 2) Transparence ; 3) Accessibilité ; 4) Performance de rapport consistante ; 5) Architecture client-serveur ; 6) Dimension générique ; 7) Traitement matriciel dynamique en fonction de l'espace ; 8) Support pour tous les usagers ; 9) Opérations de croisement dimensionnel illimitées ; 10) Manipulations intuitives des données ; 11) Rapports flexibles ; 12) Niveaux d'agrégation et dimensionnels illimités. Les points clés sont, selon Jambu, la capacité d'extraction de données, la modélisation relationnelle ou non, la capacité des programmes d'analyse de données de travailler directement sur le DW et le coût en développement, intégration et maintien.

L'étape de la préparation des données est en connexion directe avec le DW. La création des deux environnements - l'opérationnel et l'analytique - dans les organisations intelligentes, permet d'automatiser, le maximum possible, de long processus de nettoyage et préparation, en DW dénommé ETC (**Extraction, Transformation et Charge** ou **Extraction, Transformation and Loading, ETL**). Ainsi le ré-travail est évité et augmentée la productivité et la rapidité de l'analyse et de tout le processus de création de l'information nécessaire.

Dans l'ETC sont utilisés des outils spécifiques, capables d'accéder à des données directement et de les nettoyer. L'ETC exige aussi une carte des données existantes et disponibles dans l'organisation et dans les bases de données. Cette description en détail des champs des bases de données est connue comme "metadata" ou "metadonnées" dans la littérature d'informatique. Le site "**Data Warehousing and Business Intelligence**"¹ signale qu'entre les outils les plus populaires de l'ETC sont : **Data Junction, Ascential DataStage, Ab Initio** et **Informatica**.

En citant Penteado (2005, p.36)

l'orientation du DW est pour la décision, par moyen d'une interface amiable et amplement disponible. Le mot-clé est démocratisation de l'accès aux données. Pour Nobrega (2001), le DW signifie "octroyer autonomie à l'utilisateur pour qu'il obtienne lui-même ses rapports, sans avoir besoin de les commander au personnel des services d'informatique" (p.286).

Wilken (1999, apud. Cobb, 2003) mentionne que les organisations dupliquent leur quantité de données de cinq en cinq ans, "en provoquant des redondances, des inconsistances et des problèmes dans sa gestion" (p.83). Cela étant, Cobb (2003) rapporte qu'"une des fonctions importantes du DW est celle d'organiser la surabondance de données qu'existe typiquement dans la majorité des organisations"(p.83). La structure du DW permet aussi que des plusieurs personnes de différents départements ouvrent une même base de données et réalisent des analyses indépendantes. Un autre important élément du DW (Cobb, 2003, p.83) est la possibilité d'intégrer dans un seul environnement une grande variété de bases, tant de sources internes qu'externes.

Penteado et al. (2003a) indiquent que "stocker données et utiliser seulement 10 ou 20% de ses utilisations possibles conduit à une perte de temps et aussi à un affaiblissement stratégique de

l'organisation" (p.6). Ainsi, un des items essentiels pour la fonction de l'intelligence dans les organisations est l'existence de bases de données satisfaisantes, structurées de manière à faciliter et à viabiliser la réalisation d'analyses. Ces bases deviennent, par conséquent, des outils stratégiques de l'organisation.

La création d'une infrastructure spécifique avec des serveurs, de l'espace disque, des outils et des analystes pour des opérations d'analyse de données représente un avantage compétitif important. Quand elles montent ces systèmes, les entreprises font une mise à jour décisionnelle et restent un degré en dessus de la concurrence.

Préparation de données et de textes : Condition de qualité

Pereira et al. (1999) rapportent que "les bases de données sont fondamentales pour la tâche de ces qui militent dans l'analyse de Science et Technologie (S&T), en ses relations avec le développement social et économique et avec la sphère politique la plus ample en introduisant dans la discussion la question de la qualité des données qui subsidient des telles analyses" (p.4) et que doit concerner "toutes les étapes de traitement de l'information, dès sa création jusqu'à son utilisation finale" (p.4).

Gregolin et al. (2002) sont emphatiques sur le fait que les bases bibliographiques comme la Web of Science (WOS) ont été créés "pour l'identification de la littérature et non pour l'analyse de la production scientifique" (p.4). Une révision de Hood et Wilson (2003) rapporte que les études infométriques qui produisent des indicateurs à partir de bases de données bibliographiques "sont seulement une utilisation secondaire" (p.587) de ces bases. Après réviser presque 80 articles, Hood et Wilson indiquent que pour quelques-uns de ces études, l'unique solution viable est de télécharger les données, les corriger et les analyser. Les failles rencontrées "sont des erreurs ou des inconsistances des propres données, des problèmes avec l'étendue, la superposition et les changements dans les bases et, par fin, des problèmes et limitations des outils analytiques offertes par les producteurs des bases" (p.587).

Van Raan (1997, apud. Faria, 2001, p.22) affirme que telle préparation doit s'ajuster aux conditions en lesquelles le traitement est réalisé. Pour des études focalisées et d'une durée relativement courte, comme des travaux d'intelligence économique, la préparation des données normalement est développée spécifiquement pour cette tâche. Pour des études

¹ Disponible en "<http://www.1keydata.com/datawarehousing/datawarehouse.html>"

systematiques et de durée plus longue, comme la production d'indicateurs pour l'évaluation de la production scientifique, la préparation des données devient un processus continu et systématique, en visant à la création d'un système d'information pérenne, amélioré au long du temps avec l'expérience résultante d'applications successives.

Cela étant, Penteadó (2005, p.36) observe que

l'étape de traitement et préparation des données et textes influence directement la qualité et dans la fiabilité des analyses. (...) Les bases de données, dans la plupart des cas, ne sont pas développées pour des analyses. L'emphase, dans la majorité d'elles, est de permettre une récupération rapide de l'information et générer des informations opérationnelles. Dans le champ de l'intelligence économique il est commun de construire une base de données spécifique pour les analyses. Ils sont rares les cas où des données récupérées d'une base peuvent être utilisés directement par les logiciels d'analyse. (Quoniam et d'autres, 1993 ; Moguee, 1997). Plusieurs fois, la structure de l'information récupérée n'est pas compatible avec le standard nécessaire pour l'utilisation des logiciels et il est nécessaire d'incorporer des données externes. Ceci se passe parce que de telles bases ont été développées pour identifier et localiser des documents et non pour l'analyse. Alors, avant d'analyser, il est nécessaire de préparer et d'intégrer des données de différentes origines et bases, normaliser des noms, regrouper des concepts et réorganiser des champs, entre d'autres options (p.36).

En résumant, toujours il est nécessaire d'améliorer des données ou de créer des nouvelles données pour les étudier en profondeur. Il y a des outils logiciels spécifiques pour cela. Un d'eux est l'Infotrans².

Penteadó (2005, p.37) signale que

la préparation, le traitement et le nettoyage des données requièrent, généralement, la majorité du temps dépensé dans un processus d'analyse automatique de l'information. Leeds (2000) estime que dans la préparation des données il est dépensé jusqu'à 60% du temps total. Avant la préparation, 20% du temps total est utilisé pour l'identification des nécessités d'information. Après la préparation, 10% du temps total est dépensé au traitement des données et d'autres 10% en l'analyse des résultats et l'assimilation de la connaissance (p.37).

Pour Quoniam (1986), "80% du temps nécessaire pour une analyse (...) doit être dédié à la constitution de la base, à la vérification de la qualité de la base, à la reformatage des champs quand possible et au choix des méthodes d'analyse et son adéquation technique à la recherche réalisée" (p.119). Il explique qu'une bonne partie de tout ce travail est faiblement automatisée.

² De l'IUK GmbH.

Ainsi, il continue, l'activité d'analyse "est pour des spécialistes" (p.119) et celle-ci est une des raisons de l'importance de la formation de professionnels aux organisations dans ce champ.

Un différentiel dans la conversion des données

Hood et Wilson (2003, p.604) mentionnent une étude de Burton (1998) qui indique sept pas pour extraire des informations d'une base de données dans la forme appropriée pour des analyses automatiques :

1. Raffiner la stratégie de recherche jusqu'à ce que des résultats satisfaisants soient obtenus ;
2. Télécharger les références (citations) dans le format le plus ample possible et le plus explicite possible ;
3. Répéter les étapes ci dessus pour chaque base de données pertinente, en créant des archives de citations de chacune ;
4. Ajouter des champs vides ou manquants selon la nécessité de l'analyse (exemple : Année de publication, pays d'origine, langue) ;
5. Traduire les références pour un registre commun ;
6. Identifier les références (citations) dupliqués ;
7. Eliminer les références dupliquées moins complètes.

Un des outils qui s'adaptent le mieux a ces tâches listées par Hood et Wilson est le logiciel Infotrans, un outil puissant pour le nettoyage, intégration, reformatage et reconversion de bases de données en format ASCII sans limites de taille.

Le logiciel a des versions DOS/Windows et UNIX. Ses principaux attributs sont la reformatage et conversion de données ASCII entre plusieurs logiciels et systèmes. Il permet d'importer et d'exporter des bases en intégrant des bases externes à des systèmes et bases de données propriétaires d'une organisation. Il fait que les bases "conversent" entre elles et assure le transfert de toutes les données quand il y a, dans les organisations, des fusions, des incorporations, des intégrations, des mises à jour ou un changement de systèmes opérationnels ou de systèmes d'administration de bases de données. Seulement ceux qui ont déjà passé par ces situations si communes aujourd'hui dans une organisation, savent combien ceci signifie en termes de temps et d'argent. Usuellement, sont utilisés, pour faire ces conversions et transferts, les fonctions macro des logiciels de traitement de texte ou des tableaux numériques, ce qui implique une grande complexité dans l'opération en plus du temps de programmation et de maintien ou encore rien n'est fait et ces informations utiles sont vite perdues par l'inaction.

Objectif et concept basique

Infotrans assume que les archives qui seront convertis/reformatés se composent d'un ou de plusieurs références (nouvelles), que chaque registre est composé d'un ou plus champs et que les champs sont rangés dans une séquence fixe. Chaque champ doit avoir un identificateur unique de champ, mais le contenu du champ peut avoir des tailles variées. Celui est son principe basique. Un nombre très grand d'archives, de formats différents, peut être traité ainsi. L'utilisateur peut définir, comme meilleur le convenir, ce qu'il traitera comme registre, comme identificateur de champ et comme contenu de champ.

Il reformate des données extraites des bases de manière à les adapter à la lecture des logiciels de traitement automatique, normalise et intègre des données extraites de bases diverses, crée des nouveaux champs, avec des informations déjà existantes dans la base ou avec des données d'autres bases, reformate le contenu des champs pour éviter des erreurs, réunit des idées et concepts et par fin, normalise des noms d'auteurs e/ou d'institutions. Il exclut aussi des champs des informations non pertinentes pour l'analyse en question.

Caractéristiques

Selon le manuel du logiciel qui dévient un compagnon inséparable et nécessaire pour le débutant (IUK GmbH, 1995), l'Infotrans reconnaît des chaînes de caractères par le type de structure. Il permet de rencontrer et de substituer des chaînes de caractères par son texte intégral ou alors utiliser des chaînes de caractères identiques, mais composés selon la volonté du programmeur.

Celle est une de ces avantages. Les conversions peuvent être faites en tout le registre (nouvelle) ou en chaque champ d'un registre. L'objet d'une commande de conversion n'est pas un archive tout entier qu'est traité séquentiellement du début à la fin de l'archive, après chaque commande de "rencontrer et remplacer" - comme se passe avec un logiciel de traitement de texte - mais si un certain registre de la base ou même un certain champ de ce registre. Alors, les commandes de conversion sont définies à l'Infotrans avec beaucoup plus de précision et spécificité qu'aucun logiciel de traitement de texte. Ceci implique aussi en plus de rapidité et en commandes plus simples, même si le travail total de reformatage doit être divisé en plusieurs sous-pas.

Peuvent être faites aussi des substitutions entre champs. Transports de contenus d'un champ à d'autre ou de parties d'un champ à d'autre ou même pour un nouveau champ, qui sera créé et additionné au registre.

Il permet l'utilisation de listes de substitution. Le logiciel remplace des chaînes de caractères spécifiques (Ex. IN et FR) par textes préalablement classifiés en listes de correspondances (Ex. Anglais et Français). Très utile pour des abréviations, par exemple.

Tout le processus de conversion peut être suivi au moniteur. Chaque command de conversion est suivi de l'état "avant" la transformation jusqu'à l'état "après" la transformation. Ceci permet d'accompagner la conversion pas à pas, en donnant une ample transparence au processus. Quand tout fonctionne bien, en des grandes archives, il est possible d'utiliser l'Infotrans 32 bits, qu'utilise les avantages de la mémoire étendue pour traiter plus rapidement les conversions.

Autre ressource très utile est le contrôle des duplicata. En combinant des données de divers champs du registre, il est possible de créer des codes de contrôle pour identifier les duplicata, appelées **dubcodes**. Chaque code de contrôle du registre est ensuite comparé aux codes de contrôle de toutes les autres références. Les doublons ainsi identifiés, deux ou plusieurs références sont effacés automatiquement ou manuellement et tout ce processus reste très bien documenté.

La construction des dubcodes

Quoniam (1999) signale que pour construire un bon **dubcode** il est nécessaire d'effectuer beaucoup d'expérimentation pour découvrir les meilleures clés. C'est un jeu éternel entre trouver "beaucoup de bruit", un **dubcode** laisse dans la base beaucoup de doublons et "silence", quand un **dubcode** retire de la base des références qui ne sont pas des doublons. La règle générale est plus le **dubcode** est petit, plus grand sera le nombre de références considérées égales. Ainsi, la plus grande charge de travail devient de trouver une manière de construire le **dubcode** en réduisant sa taille mais en conservant son efficience pour identifier les références doublées.

Le manuel de l'Infotrans (IUK, 1995, pp.1-16) insiste que le plus grand problème dans le contrôle des doublons n'est pas "la reconnaissance de références qui concordent entièrement mais, en son lieu, reconnaître les références qui sont seulement similaires mais que

représentent la même publication" (p.2). Deux références que sont plus au moins différentes mais qui font référence à un même objet doivent être attribuées

*à un groupe de doublons. Si ceci n'e se passe pas, le résultat est une **non-identification**. Deux ou plus références que diffèrent plus ou moins marquéement un de l'autre mais qui représentent deux ou plus objets doivent être maintenus séparés. Si ceci ne se passe pas, le résultat est une **identification incorrecte** (p.2).*

L'IUK (1995) considère que

*les **dubcodes** pour entreprises et adresses doivent être structurées d'une manière différente des **dubcodes** pour des bases de données de textes ou de patents (...) doivent être analysés la structure des références et les contenus des champs de la base de données dans laquelle les objets sont présentés (...) et déterminer, pour chaque base de données, un groupe de références déterminé parmi lesquels les éléments et données seront appropriés pour une claire identification des objets (p.2).*

Elle recommande que seulement les éléments des données qu'identifient clairement une caractéristique, laquelle est espérée, sera présente aussi dans les références qui seront comparés, doivent être choisis pour le **dubcode**. Cela étant, les meilleurs identificateurs d'un travail écrit sont son auteur et son titre. Autres items importants sont la date et le lieu de la publication et d'autres "éléments que présentent les caractéristiques citées de façon codifiée" (p.2), entre eux, l'ISBN ou l'ISSN. La manière suggérée pour créer le **dubcode** est en sélectionnant deux à quatre lettres des premières lettres des noms des auteurs, des mots du titre, deux derniers chiffres de l'année de publication et ainsi de suite. De la même manière, l'IUK (1995) soutient que,

des segments de registre présents en toutes les références mais en lesquels l'expérience ait montré qu'ils varient en fonction de différences ou de normes d'indexation appliquées de manière inconsistante doivent être omises parce qu'ils vont résulter dans le non-identification des doublons (p3).

Analyse et synthèse de l'information

Boutin (1999) indique que "l'accroissement tant quantitatif que qualitatif de l'information disponible" (p.16), conduit à un sérieux déséquilibre entre l'offre et la demande d'information. D'un côté, beaucoup d'informations massives, brutes, dispersées et non-validées sont offertes. De l'autre côté, ce que les organisations veulent est une information spécifique, synthétique, validée et de facile accès.

À cause de cette dysfonction croissante entre offre et demande d'information, Boutin (1999, pp.17-18) recommande qu'il est "nécessaire d'en faire une révolution dans la gestion de l'information" (p.17), au côté du développement de nouvelles compétences dans les organisations :

1. Devant la multiplicité des sources d'information disponibles, il est nécessaire de savoir où trouver une information spécifique ;
2. Devant le volume des informations disponibles, il est nécessaire être capable d'extraire d'une grande masse d'informations, une information pertinente à forte valeur ajoutée (p.18).

Cela étant, Dutheil (1991, apud. Boutin, 1999) signale que "le pouvoir n'appartient plus à celui que détienne l'information mais si à celui qui sait la traiter" (p.18). Boutin (1999) observe que traitement d'information et prise de décision sont "deux dimensions indissociables" (p.18) et que le traitement seulement existe et est pertinent parce qu'il mène à une action et cette décision seulement est possible parce qu'elle a comme base un traitement. L'information analysée et validée à pourtant une grande valeur parce qu'elle "amène une solution à un problème défini" (p.18). C'est comme une pépite d'or trouvée après le traitement d'une énorme masse de minerai.

Cobb (2003) rapporte que la majorité des informations desquelles une organisation a besoin "est dedans elle-même, surtout avec ses employés. Le problème consiste à savoir réunir et synthétiser cette connaissance en un format employable et qui bénéficie la planification et les opérations de l'organisation" (p.80). Cobb indique qu'une solution consiste à développer l'habileté d'analyser des informations publiques pour combler ces "vides de connaissances" (p.80).

En faisant référence à l'information technique, Rouach (1999) explique qu'en fonction de la concurrence, tous les pays cherchent à "dominer et tirer profit" (p.49) de ce type d'information, "aussi tant que la **gestion stratégique de l'information est devenue un des moteurs de la performance globale, des entreprises et des nations**" (p.49).

Les techniques et les méthodes d'analyse concernant des bases de données sont diverses, depuis les bases formelles et structurées aux bases non-structurées, créées spécialement pour l'analyse à réaliser. Ces techniques et méthodes peuvent être divisées en deux versantes : L'analyse de données et l'analyse de textes.

L'analyse et l'exploration de données

Penteado (2005, p.37) rapporte que

*selon Jambu (2000, p.8), le processus de l'exploration de données (**datamining-DM**) combine plusieurs méthodes mathématiques, statistiques ou ayant origine dans des algorithmes, pour déterminer une solution à un problème, dans un univers décisionnel. Porter (2003, p.3) indique qu'en général, quand nous travaillons avec les techniques de **datamining**, nous extrayons des informations utiles de n'importe quelle type de données. Mais le plus commun est d'utiliser de données numériques et, pourtant, quantitatives. Une bonne connaissance de Statistique est nécessaire. Cette caractéristique générale influence le choix des outils de **datamining**, aussi orientées pour le traitement quantitatif (p.37).*

Parmi ces outils nous pouvons citer des logiciels de statistique qui ont développé des suites spécifiques (SAS³, SPSS⁴ et Statistica⁵) et des logiciels libres R⁶, de statistique et le Weka⁷, spécialisé dans l'exploration de données.

Le cycle d'analyse dans l'exploration de données

Jambu a nommé 16 étapes pour le cycle d'analyse et de gestion en exploration de données, très semblables au cycle de l'intelligence cité par Faria et Quoniam (2002) :

1. Définition de l'objet de l'étude ;
2. Définition des objectifs à atteindre ;
3. Étude des données existantes ;

Il est considéré que les données doivent être liées à l'objectif et traduire des faits précis en relation à l'objectif. L'analyste doit définir les variables, les entités et les périodes de temps et il a toujours deux difficultés primordiales : La donnée supplémentaire considérée comme erreur et la donnée manquante, considérée comme lacune.

4. Création de données spécialisées/ou de bases externes/particulières pour l'analyse ;
5. Définition des traitements spécialisés nécessaires ;
6. Collecte des données ;

³ Informations en "[http : //www.sas.com/](http://www.sas.com/)"

⁴ Informations en "[http : //www.spss.com/](http://www.spss.com/)"

⁵ Informations en "[http : //www.statsoft.com/index.htm](http://www.statsoft.com/index.htm)"

⁶ Informations en "[http : //www.r-project.org/](http://www.r-project.org/)"

⁷ Informations en "[http : //www.cs.waikato.ac.nz/~ml/weka/](http://www.cs.waikato.ac.nz/~ml/weka/)"

7. Traitement ou importation des données ;
8. Contrôle des données ;
9. Gestion des données ;
10. Accès aux données par les utilisateurs ;
11. Analyse interactive des données (par l'utilisateur lui-même) ;
12. Exploration des données :

Un processus itératif, parfois, interactif, qui combine plusieurs méthodes mathématiques, statistiques ou algorithmiques pour déterminer une solution à un problème, dans un univers décisionnel (Jambu, p.50) ;

13. Elaboration des tableaux et tables d'accompagnement, des rapports (titre, objectif, bref description des données, liste des tableaux, graphes, résultats et, pour chaque résultat, sa relation avec l'objectif, légendes et citation des sources) ;
14. Décisions ;
15. Plans d'action ;
16. Action.

Pereira (2005) rappelle qu'en explorant "des grandes quantités de données" (p.8) on peut identifier les relations systématiques ou normatives entre les variables et ensuite les valider "en appliquant les normes détectées à des nouveaux sous-ensembles de données. Cette étape a trois phases : **Exploitation ; Construction du modèle ou définition des normes ; Validation/vérification**" (p.8).

Penteado (2005, pp.37-38) présente les cinq façons utilisées par Sulaiman et Souza (2001, p.267) pour "générer des résultats par l'exploration des données (**datamining**)" : Des règles d'association ; Des hiérarchies de classification ; Des normes séquentielles, normes de séries temporelles ; Catégorisation et segmentation, que nous analysons ensuite :

1. Des règles d'association. Elles visent à "rencontrer des items en une transaction qui puissent déterminer la présence d'autres items dans la même transaction" (p.267).
Exemple : Celui qu'achète du pain et du lait, achète aussi du beurre ;
2. Des hiérarchies de classification. Elles créent "un modèle basé sur les données connus" (p.267) et aident à expliquer le pourquoi d'une classification donnée et aussi permettent de "classer de nouvelles données à partir d'une classification existante" (p.267). Exemple :

Créer des limites pour la concession de crédit basés sur l'historique des transactions de crédit antérieures ;

3. Normes séquentielles. Elles indiquent des comportements ou une séquence de comportements. Exemple : "Toutes les fois qu'une jeune femme achète des chaussures en cuir, elle achètera aussi des sacs à main et des ceintures dans les prochains trente jours" (p.268) ;
4. Normes de séries temporelles. Elles montrent des occurrences similaires dans un espace de temps. Aux données en dessus, on additionne la saison : À l'automne, les jeunes femmes achètent des chaussures, des sacs à main et des ceintures. Au printemps ce schéma change pour des sandales, des sacs à main et des chapeaux ;
5. Catégorisation et segmentation. Des références avec des caractéristiques similaires sont réunies. Exemple : Un groupe de consommateurs peut être classé comme "peu acheteur", "moyennement acheteur" ou "très acheteur" (p.268) pour un produit déterminé (pp. 37-38).

Penteado (2005, p.38) poursuit :

Dans la plupart des cas, nous travaillons dans l'exploration de données en cherchant à identifier les profils des différents sujets d'analyse (des usagers et/ou des clients) et ses diverses normes de comportement ou de consommation. Ensuite, en travaillant par segments, il est possible d'identifier le profil des meilleurs clients ; Les produits et services consommés par ou les caractéristiques de chaque segment de clients ; Les normes de consommation des produits, comment et quand deux ou plus produits ou caractéristiques s'associent ou se réunissent dans un seul achat ou événement et aussi les normes de consommation/comportement par région, arrondissement, age ou sexe (p.38).

Pour finir,

*Jambu (2000, p.93) rapporte les champs d'applications typiques de l'exploration des données : Les études de cas, l'amélioration de la qualité et de l'efficacité du négoce, la satisfaction des employés, la relation avec les clients, le **marketing** dirigée, le **marketing** local, les indicateurs de négoce, la prévision, les tendances, la veille de la concurrence, des achats, de la technologie et de l'information, le contrôle des dépenses, les finances et les mouvements financiers et les processus de gestion (p.38).*

Exemples d'applications intelligentes de l'exploration de données

Des couches et de la bière

Penteado (2005, p.38) rapporte un cas classique d'exploration de données :

*Il a été découvert que le profil du consommateur de bière était similaire à celui du consommateur de couches : Des hommes mariés, avec 25 à 30 ans, qu'achetaient les deux produits les vendredis, période de fin d'après-midi, début de la nuit. La **Wal-Mart** - chaîne de magasins américaine - a décidé alors mettre les gondoles*

de couches au côté de celles de bière. Résultat : Les ventes des deux produits ont augmenté de 30% les vendredis.

Des bonnes relations avec le client

La **Telemar**, une des plus grandes entreprises de télécommunication de l'Amérique Latine, avec une couverture de 64% du territoire brésilien, en 16 des 27 états du Brésil (départements), a intégré sa base de données de presque 25 millions de clients (Balaj, 2005a) dans un entrepôt de données. Selon Balaj (2005b, pp.24-27), le système a permis augmenter les ventes, réduire les débranchements et réduire des coûts. Entre 2003 et 2005, la participation de l'entreprise dans le marché d'Internet à haute vitesse est passé de 7% à 25%, soit un accroissement de 280%. La **Telemar** a été capable "d'anticiper le potentiel de croissance du marché de portables, en maintenant l'expansion de sa base de clients" (p.26). La part des nouveaux services (portable, transmission de données, Internet haute vélocité et liaisons entre les états), en opposition aux services traditionnels (téléphone, réseaux et téléphones publiques), dans le total des recettes a augmenté de 30% en 2003 de 39% en 2005, même avec une croissance réelle des recettes, en ce dernier an, de 9% (p.27).

Crédit et insolvabilité

Des modèles mathématiques sont créés pour calculer les probabilités d'opérations financières et alimentés avec les variables qu'expliquent l'insolvabilité. Ces modèles sont traités dans les bases de clients des institutions financières. Pereira (2005, pp.15-18) signale que par ce moyen sont identifiés des clients "ayant grande chance de ne pas payer" (p.125) évitant ainsi la concession de crédits de difficile récupération. Une simulation a été faite avec une liste de 252 clients insolubles d'une institution financière et a identifié 196 (78%) comme des mauvais payeurs en potentiel.

Générer plus de négoce

Le **Groupe Accor** intègre au Brésil 29 marques entre des hôtels, des agences de voyages et des services divers. Les données des clients de tout le groupe ont été concentrées dans un entrepôt de données appelé Table de Négoces. Son objectif, explique Angerame (2005) est de "garantir une vision unique du client corporatif Accor, promouvoir l'intelligence de négoce, des façons des alliances et des synergies entre les marques" (p.7) et favoriser une action intégrée du groupe, développer plus les nouveaux négoce, fidéliser des clients, générer plus

de profit. Angerame observe qu'une analyse intégrant les données disponibles sur le marché hôtelier de la Grande Sao Paulo a permis d'identifier l'impact du commerce et de la concurrence et des régions pour des nouvelles affaires, en générant encore une vision globale de ce marché et en signalisant 2500 entreprises pour prospector des nouveaux négoce avec le groupe. Des analyses de synergie de négoce intra et extra groupe Accor ont indiqué, par exemple, des chemins pour l'expansion des réseaux d'affaires du groupe soit avec des grandes entreprises, soit avec des secteurs entiers et aussi d'accompagner de près ses 100 plus grands clients (pp.17-36).

Veiller le marché

Le **Groupe Carrefour** a un entrepôt de données où sont intégrées des données sous dix perspectives : Lieu de ventes, public-cible, produits, prix, promotion, présentation (dessin), publicité, positionnement, personnes et planification. Inafuco (2005, pp.14-17) affirme que le système d'intelligence de marché génère entre d'autres produits stratégiques, des études du potentiel du marché, de l'impacte des nouveaux magasins, de fusions et d'acquisitions, de profil des clients, des tendances du consommateur, dépenses moyennes, d'offres stratégiques, de compétitivité des prix, d'impacts des promotions et des recherches d'image et de qualité de service.

Des contrats de maintenance

La **Sears Roebuck** a un système qui veille ses ventes d'électroménager - qui aux États-Unis sont vendus en conjonction avec des contrats de maintien. Le système informe les clients par lettres quand est l'époque de renouveler les contrats, ce qui maintient une importante source de recettes et améliore aussi la qualité du service et la satisfaction des clients (Reynolds, 1992, p.223). Des systèmes similaires sont utilisés partout dans le monde. À Brasilia, un magasin, le **Mundo dos Filtros**, a un système pour la maintenance de filtres d'eau douce dans les résidences opère par téléphone.

Facteurs critiques de succès (FCS)

La **Kubota Tractor**, entreprise de pièces détachées de ce type de véhicule considère un FCS, minimiser le temps qu'une pièce reste au dépôt, ce qui permet de maintenir bas les coûts de stockage des produits et fait qui les pièces soient délivrées rapidement à leurs partenaires. Des systèmes d'intelligence ont été créés. Ils permettent, d'entre autres fonctions, de comparer la

demande de pièces avec la demande projetée et d'anticiper les réquisitions de pièces sans lever trop le stock et aider aussi à décider sur des quantités à demander en évitant la faute de pièces en stock (Reynolds, 1992, p.240).

L'analyse et l'exploration de textes

Penteado (2005, p.39) indique que

*l'analyse et l'exploration de textes (**text mining**) s'est développée à partir des travaux de quelques précurseurs comme Derek de Solla Price, Henry Small, Antony van Raan, Donald Swanson, Henri Dou et Alan Porter, en textes structurés, en opposition à des textes non-structurés ou libres. C'est-à-dire, utiliser des bases de données internes ou externes, comme la Dialog⁸, Pascal et Francis⁹, avec une structure de références en indiquant, par exemple, "auteur", "titre", "date de publication", "mots-clés" et extraire des informations sur l'évolution et le développement des activités scientifiques, gestion, évaluation et productivité des activités de C&T, construire des indicateurs (citation), des réseaux (communautés) et veiller sur des innovations, des technologies et des concurrents (p.39).*

"Les travaux d'exploitation et d'analyse de textes s'inscrivent dans des champs comme infométrie, bibliométrie, scientométrie, médiométrie, museométrie et webométrie", continue Penteado (2005, p.40). Le Coadic (2003) mentionne que l'infométrie est "un nouveau champ de recherches en science de l'information" (p.2). Il ; Ajoute :

À l'intérieur de l'infométrie sont regroupés les sous-champs de recherches formés sur des secteurs informationnels spécialisés comme celui du livre, la bibliométrie (la première née), de la R&D (recherche-développement), la scientométrie, des mass-médias, la médiométrie, des musées, museométrie et du WorldWideWeb, la webométrie (la dernière-née) (p.2).

Quoniam (1996, p.31) indique que "l'ensemble de ces méthodes peuvent être regroupés sous le terme **statistique textuelle**". Il cite encore les livres "Analyse statistique des données textuelles" (1988) et "Statistique textuelle" (1994), de Ludovic Labart et André Salem, édités chez Dunod, comme les "ouvrages généraux" de ce thème.

⁸ Le système Dialog (<http://www.dialog.com/>), du groupe Thomson, constitue aujourd'hui le plus grand serveur de données du monde. Basé en North Caroline (EUA) il a été créé en 1966 par Roger Summit. Il a offert le premier la récupération d'informations en ligne avec un système internationale. Le Dialog réunit plus de 900 bases avec plus de 1,4 milliards de références, presque 12 Terabytes d'information. Une bonne partie de la connaissance mondiale est ainsi disponible pour 20 millions de clients dans une centaine de pays.

Les distributions des mots ont des propriétés et caractéristiques particulières et elles doivent être considérées, selon le type de problème et de données travaillées, pour choisir le traitement adéquat. Ces distributions et leurs analyses font référence à des différents aspects de l'information et de sa qualité. Leur principal matière-première est le mot. Il peut représenter, d'entre autres choses, un concept ou un thème, un individu, une organisation ou encore un ensemble de thèmes, d'individus ou d'organisations.

Le Coadic (2003) explique que "l'étude des phénomènes informationnels a révélé l'existence de régularités, de rapports mesurables, de distributions" (p.2) qui seules peuvent être actualisées avec l'application de la mathématique et de la statistique. Entre ces "régularités", Lafouge et al. (2003) mentionnent la loi énoncée par Alfred Lotka, en 1926, "qu'établit une relation simple entre le numéro d'auteurs et le numéro d'articles publiés par ces auteurs" (p.17) ; La méthode dévoilée par Samuel Bradford, en 1934, qui permet de "déterminer comment les articles publiés sur un sujet sont distribués dans les revues scientifiques et techniques" (p.17) ; La loi que détermine la "fréquence d'apparition des mots dans un texte" (p.17) révélée en 1935 par George Zipf ; Une théorie statistique qu'explique "la probabilité de transmission d'un signal électrique en une ligne de transmission" (p.18) élaborée en 1949 par Claude Shannon ; L'application de "modèles de prévision retirés de la recherche opérationnelle pour la gestion de la circulation des oeuvres dans les bibliothèques" (p.18) idéalisée par Philip Morse en 1968 ; Et le modèle probabiliste pour expliquer "des divers phénomènes caractéristiques des productions bibliométriques" (p.18) proposé par Derek de Solla Price en 1976.

En appliquant à ces phénomènes la mathématique et la statistique, les scientifiques de l'information peuvent mieux décrire, expliquer et faire des prévisions sur les faits informationnels. Le Coadic signale que cette plus grande capacité explicative des phénomènes inclut de plus "une capacité unificatrice (expliquer, c'est ramener la diversité des phénomènes à un très petit nombre de principes) et une capacité générative (suggérer des concepts nouveaux, des stratégies nouvelles)" (p.2). Néanmoins Le Coadic (2003) alerte que quand nous utilisons les techniques mathématiques "sur n'importe quoi et n'importe comment" (Xirdal, 1976, apud. Le Coadic, 2003, p.3), il y a des risques que les analyses et conclusions deviennent non fiables.

⁹ La Pascal est une base de données multidisciplinaire et multilingue. Elle couvre les Sciences, les Technologies et la Médecine. La base Francis couvre les Sciences Humaines et Sociales. Les deux sont produites par l'Institut de l'Information Scientifique et Technique de la France (INIST CNRS).

Une particularité importante est que les distributions rencontrées dans des textes, dites "zipfiennes" (lois de Zipf, Lotka et Bradford¹⁰) exigent un traitement spécifique, différent de la statistique conventionnelle qui, d'une manière générale, traite plus spécifiquement des distributions normales ou assimilées.

En conséquence, Penteadó (2005, p.40) rapporte que

plusieurs logiciels ont surgi pour accomplir des fonctions spécifiques de cette analyse des distributions zipfiennes en traitant de l'exploration, le positionnement, la structuration et la prospection de textes structurés. Une partie des logiciels est apparue aux années 90 et sont bâtis sur le système opérationnel DOS (p.40).

En version Windows trois logiciels se détachent : Matheo Analyser¹¹, VantagePoint¹² et WinIDAMS¹³. Les deux premiers sont des produits d'équipes de chercheurs dirigés par Henri Dou, à l'Université d'Aix Marseille III et par Alan Porter, à l'Université de Georgia Tech. Ils permettent l'intégration des fonctions de traitement bibliométrique, traitement statistique et la représentation graphique dans un seul environnement. Le troisième logiciel est un hybride entre un logiciel de statistique et de bibliométrie. Il fait partie d'une famille de logiciels développés par l'Unesco et distribuée gratuitement, Isis-ASCII-WinISIS-WinIDIS-WinIDAMS-GenIsisWeb, qui permet, respectivement, la création, l'importation, l'analyse de bases de données et de textes et leur disponibilité sur l'Internet. À mentionner encore deux logiciels précurseurs développés à l'Université d'Aix Marseille III, Dataview et Matrisme. Ils ont comme principe basic la construction de groupes par propagation et la représentation des relations entre les éléments constituant d'un groupe par moyen d'un réseau (Boutin, 2001 ; Quoniam, 2001).

¹⁰ Rostaing (1993, pp.74-91) relate d'une manière très complète les caractéristiques de ces trois lois.

¹¹ La famille du Matheo Analyser, disponible sur le marché depuis 2003, inclut aussi le Matheo Patent pour des analyses spécifiques en bases de brevets (dans le cas, la base de brevets EPO - **European Patent Office**, accessible gratuitement sur l'Internet). Les brevets réunissent des informations uniques, non rencontrées en aucune autre place. Par exemple : a) quelles entreprises réalisent des travaux de pointe; quelles sont les leaders; b) quels individus réalisent des travaux de pointe; c) quels pays sont leaders d'une technologie; d) quant temps les entreprises prennent pour utiliser un brevet; combien de temps de R&D elles nécessitent pour se transformer en profit; e) qu'elles technologies sont en accroissement ou en décadence; où l'argent de R&D est appliqué entre les leaders d'une industrie et f) les relations entre des entreprises liées à des R&D similaires ou en produisant les mêmes produits; les relations de recherche entre des entreprises subsidiaires.

¹² Le VantagePoint est disponible sur le marché depuis 2000, par Search Technology Inc.

¹³ Le WinIDAMS de l'Unesco a des versions en anglais, français et espagnol. Il est distribué gratuitement depuis 1988.

La méthode d'analyse bibliométrique

Quoniam (1996, pp.10-34) décrit en détail la démarche d'analyse de textes, dénommée analyse bibliométrique. Le processus commence par la segmentation du texte en **formes graphiques**. Ensuite, il faut choisir des **caractères délimiteurs** qui définiront par exclusion les autres caractères considérés "**comme caractères non-délimiteurs**". Une suite de caractères non-délimiteurs bornés à ses deux extrémités par de caractères délimiteurs est une **occurrence**. Deux suites identiques de caractères non-délimiteurs constituent deux occurrences d'une même **forme graphique**" (p.34).

Les prochains pas sont la transcription des données "sous la forme de tableau logique (de présence/absence de mots" (p.10) (ou concepts) dans le corpus en analyse et après, ou encore, sous forme de tableau de "contingence croisant deux ou plusieurs concepts (auteurs et mots-clés, par exemple)" (p.10).

D'après Quoniam en travaillant avec des mots, "la distribution de fréquences s'effectue selon la loi de Zipf, caractéristique de diversité (des classifications automatiques non hiérarchiques avec autodétermination du nombre de classes)" (p.11). On suit le même schème général des applications statistiques : Définir le problème, obtenir les données, traiter les données et les interpréter. Quoniam (1996) rapporte que **la méthode part du problème à résoudre**. Quand il s'agit d'analyses stratégiques "la question est dictée par la stratégie" (p.33) et, communément, "est énoncée par des tiers" (p.33). Cela étant, "la définition du problème est une phase cruciale dans le processus" (p.34). La principale caractéristique de cette phase doit être son itérativité. Un autre point est qu'en fonction de la diversité des sujets, il n'est pas possible de travailler seul. L'analyse est faite "toujours en collaboration avec les spécialistes du champ étudié" (p.34).

Pour Jakobiak (1991) **l'information existe "mais il est nécessaire savoir où la trouver"** (p.148). Il explique que "le véritable problème est aller trouver les informations difficiles à trouver" (p.148). Quoniam (1996, p.35) renforce ce point en argumentant que quand le travail concerne des textes structurés en bases bibliographiques publiques ou privées, il est nécessaire de savoir d'abord qu'il existe une base de données où cette information pourra être obtenue. Le choix peut ne pas être facile. Seulement au Dialog il y a plus de 900 bases. Ensuite, il faut connaître comment la base analysée a été construite, sa structure et contenu informationnel. L'analyste doit aussi être capable d'extraire de la base choisie la partie qui sera analysée.

Quoniam (1996) mentionne qu'il y a une "dépendance entre qualité des données et qualité de l'analyse" (p.36). De cette façon, "il est nécessaire d'avoir une bonne stratégie de recherche de l'information qui peut être affinée éventuellement " (p.36). Selon le résultat de la recherche, le problème peut être renégocié ou encore, la base reconstruite, par sa conversion de base bibliographique en base bibliométrique.

Les caractéristiques essentielles des informations traitées

Des **caractéristiques essentielles des informations** "qui conditionnent certains types de traitement" (p.36) sont objet de l'attention de Quoniam (1996, pp.36-40) :

1. **L'information** (l'accès aux bases de données internationales, l'infrastructure de personnel et matériel) **coûte cher**¹⁴.
2. **Dans l'analyse de données textuelles il n'a pas une maîtrise de l'échantillon.**

Selon Quoniam, quand nous avons une analyse statistique,

même si l'échantillon représentatif est sélectionné aléatoirement, la population est choisie avec soin (si l'état de satisfaction d'une catégorie est étudié, l'échantillon aléatoire sera extrait dans cette catégorie) (p.37).

Déjà en recherche bibliométrique l'échantillon ne sera pas si homogène. Seulement le producteur de la base de données bibliographiques a le total contrôle de l'information. La seule cohérence de l'échantillon bibliographique sera "la présence dans chaque référence des concepts qu'ont servi pour déterminer l'objet de recherche" (p.37).

Cela étant, rapporte Quoniam (1996), il est toujours nécessaire de vérifier, avant l'analyse, la qualité de l'information de la base. Une des manières les plus simples de faire cette vérification, c'est celle avec la construction d'un indicateur de champs vides. Un champ où cet indicateur est très commun devra être déconsidéré ou être analysé avec plus de soin. Quoniam indique aussi que "l'utilisateur devra toujours placer l'analyse dans le contexte de la base" (p.10) et, si possible, combiner plusieurs bases dans une même analyse. Il décrit un critère pour décider quand mélanger les bases et quand les analyser chacune séparément : Si la jonction implique en réduction de l'information, le mieux est d'analyser les bases séparées. S'il n'y a pas des pertes d'information, les bases peuvent être combinées.

¹⁴ Voir plus détails en "Dimension économique de l'intelligence".

3. Propriétés particulières de l'information traitée.

- a. Approche unidimensionnelle et
- b. Approche multidimensionnelle et relationnelle.

Penteado (2005, p.40) rapporte que les méthodes d'analyse des distributions zipfiennes

impliquent d'y utiliser des statistiques unidimensionnelles (combien sont-ils et que signifient les valeurs/mots), statistiques bidimensionnelles (comment sont-ils et combien mesure la relation entre deux valeurs/mots), statistiques multidimensionnelles (comment sont-ils et combien mesurent les relations entre plusieurs variables/mots) et statistiques probabilistes (détecter des comportements émergents ou atypiques ou encore comment se comporteront ses variables/mots) (p.40).

Quoniam rappelle l'attention sur deux phénomènes présents dans les bases textuelles : Les données non-mutuellement exclusives et la production de matrices creuses, avec un faible présence de descripteurs.

En enquête, lors de la constitution des questions fermées ou de leurs saisies, on cherche à rendre les modalités mutuellement exclusives. L'exemple le plus simple est constitué par les réponses de type "oui-non" : L'enquêté donne l'attribut oui ou l'attribut non, jamais les deux à la fois. Dit autrement [dans une enquête] l'espace des modalités est orthogonal. Dans les bases de données bibliographiques, l'auteur d'un laboratoire donné n'exclut pas les autres auteurs de ce même laboratoire. Il en est de même pour les affiliations, les mots-clés. Seul le journal de publication, la date de publication et quelques autres informations respectent le principe d'orthogonalité (p.38).

Et continue :

*Ce phénomène oblige à s'intéresser aux relations entre les **formes graphiques** dans chacun des champs (**intrachamp**) avant de s'intéresser aux relations entre les **formes graphiques** d'un champ à l'autre (**interchamp**).*

- c. Variété des formes graphiques.

Pour Quoniam, des différentes bases de données ont des différents niveaux de qualité. Ainsi, quand sont réunies des données provenant de diverses bases ou même d'une seule base, de périodes divers (décades différentes, par exemple), une expression de même sens peut avoir plusieurs **formes graphiques**. Il donne comme exemple les diverses variantes des noms d'un même auteur, des erreurs d'indexation et de dactylographie et les mots-clés au singulier et au pluriel¹⁵. En conséquence Quoniam considère "indispensable" (p.38) la préparation des

¹⁵ Voir plus détails en "La qualité en base de données".

données¹⁶ parce que des erreurs "auront influence sur les traitements" (p.38) et aussi sur les résultats.

d. Grande masse de données.

Quand nous décomposons un texte ou une base de données textuelle en **formes graphiques** une quantité très grande de données est créée. Quoniam recommande alors de "réduire le plus possible la taille ou l'espace du problème à traiter" (p.39).

e. Adéquation problème - base de données.

Il est important de "connaître parfaitement la structure et le contenu de la base de données variant avec chaque nouvelle base analysée" (p.39).

4. Propriétés du traitement réalisé.

a. Type de représentation désirée.

Quoniam (1996) observe, à ce point, l'opportunité de marier les préférences du destinataire de l'analyse avec le type de représentation du problème. Il souligne que "les éditions de résultats sur papier sont frustrantes par rapport à celles présentées directement sur support informatique" et que "mille références imprimées sur papier sont inutilisables" (p.39).

Quoniam signale qu'à l'occasion d'une présentation sous forme papier

les dénombrements constituent la méthode d'analyse la moins susceptible de polémiques. En effet les méthodes de positionnement ou de classification peuvent d'avantage être paramétrées pour produire des présentations orientées (p.42)

b. La simplicité des analyses.

La méthode scientifique indique que la meilleure analyse est toujours la plus simple. Mais il faut prendre garde de ne pas confondre simplicité et facilité. Il ajoute

tous ces traitements sont là pour être interprétés, transmis et commentés. (...) Une très bonne analyse mais ne pouvant pas être transmise car trop compliquée, ne sert à rien (p.40).

c. L'interprétation.

Quoniam (1996) rappelle que "parfois, il est meilleur de signaler l'impossibilité de répondre à une question que de la répondre mal" (p.36).

¹⁶ Voir plus détails en "Préparation de données et textes : condition de qualité".

La meilleure interprétation de l'information est toujours celle qui fait l'objet d'un consensus entre les différents spécialistes participant de l'analyse. Quoniam (1996) signale que l'interprétation

ne doit pas se limiter à une simple interprétation "statistique" mais elle doit répondre à la question posée en confirmant ou infirmant l'hypothèse. Si la première partie de l'interprétation (la phase purement statistique) est assez aisée, la deuxième partie [répondre à la question, confirmer ou non l'hypothèse] est la plus difficile de l'analyse parce qu'il n'y a pas ou si peu de méthodes. C'est souvent une grande part d'intuition et de concertation (p.40).

L'analyse est une phase critique, où les informations sont transformées en intelligence. Les informations collectées, beaucoup d'entre elles sans une connexion apparente sont ainsi transformées par les activités d'analyse. Les réseaux de spécialistes du champ vérifient la consistance des informations, établissent des relations et jugent leur impact pour l'organisation.

Les nouvelles frontières

Penteado (2005, pp.46-47) rapporte qu'

entre les champs de frontière de l'analyse et d'exploration de données est la méthode de découverte de connaissance en bases de données de Don Swanson (2001). Le départ est d'une observation, par exemple (Porter, 2003), les grenouilles meurent. Ensuite, les causes du déclin de la population de grenouilles sont recherchées en une base de données. Dans la troisième étape, des vétérinaires et chercheurs sont consultés si les associations trouvées et non mentionnées par la littérature ont du potentiel pour réduire la mortalité des grenouilles. Si quelqu'une association est validée par les spécialistes, une connaissance inédite aurait été découverte dans une base de données. En utilisant cette méthode, Swanson a identifié que l'huile de poisson a des effets thérapeutiques positifs pour le traitement de la Maladie de Raynaud et que la déficience en Magnésium peut donner lieu à des migraines.

Textes semi-structurés

L'analyse et l'exploration de textes semis et non structurés sont un autre champ d'avant-garde. La majorité des outils actuellement disponibles et en développement est tournée vers l'Internet et fonctionne selon le principe selon lequel le chercheur définit les mots qui seront objet de la recherche et alors des documents ou parties de documents sont récupérés dans des sites et des cartes de collaborations sont créés.

Une initiative ambitieuse est celle d'IBM qui a développé une plate-forme pour exploiter et extraire des tendances, des normes et des réseaux de collaboration de la Web : **WebFountain**. En somme, trois composants : a) Superordinateurs pour opérer le service, b) Infrastructure de stockage capable de garder **toute** l'Internet et c) Des outils d'analyse qu'incluent le traitement de langage naturel, statistique, probabilités, apprentissage de la machine, reconnaissance de tendances et intelligence artificielle.

La qualité en bases de données

Le Programme brésilien de qualité au service public (QPAP) du **Ministério do Planejamento, Orçamento e Gestão** [Ministère du Plan, Budget et Gestion] (MPO), a établi qu' "information et analyse" est un des sept items à contrôler dans la gestion publique. Ces sept critères sont : Leadership, planification stratégique, focus client, information et analyse, gestion de personnes, gestion de processus et résultats de l'organisation. Ils sont aussi évalués en dizaines de pays des cinq continents et par un des plus prestigieux concours d'excellence des États-Unis, le Prix Malcolm Baldrige.

Selon le MPO (2000) :

le succès des organisations publiques et privées dépend de leur capacité d'identifier et de traiter les informations pertinentes des environnements externe et interne, en promouvant leur organisation leur intégration et en rendant ces informations disponibles aux agents internes et externes de façon opportune et adéquate, en les transformant dans des outils du processus de prise de décision (p.91).

La convergence et l'intégration entre les bases de données et les systèmes d'informations existants dans l'organisation sont d'autres items évalués. Le QPAP considère qu'il est

fondamental que les mécanismes de sélection, collecte, enregistrement et stockage des informations soient fiables et sûrs et, principalement, que son utilisation par ses usagers soit opportune. L'information adéquatement stockée, consistante et intègre ne sert à rien, si l'administrateur public n'a pas des facilités pour y accéder au moment où il en a besoin (p.92).

Le MPO (2000) continue :

pour cela, il est important que les bases de données d'informations de diverses natures soient intégrées et convergentes, rendant possible à des systèmes différents d'échanger des informations entre-eux, en permettant le croisement de données stockées dans les diverses bases d'information (p.92).

Pereira et al. (1999) classifient "la question de la qualité comme un des aspects les plus importants de l'industrie de l'information en nos jours" (p.2). Pour eux,

l'importance des bases de données, principalement les bibliographiques, est définie par son utilisation croissante dans la production d'indicateurs de Science et Technologie (S&T) et, plus récemment, pour produire des études stratégiques de S&T, champ de la connaissance que s'organise sous la dénomination d'intelligence compétitive. Une autre utilisation est la génération de cartes de connaissance, que permettent de situer graphiquement des champs/thèmes de recherche en ses corrélations cognitives, en exhibant encore les chercheurs et les institutions orientées pour la production de ces savoirs (p.2).

Hood et Wilson (2003) signalent divers types d'erreurs dans les bases de données et affirment que si ses erreurs ne sont pas considérées et adéquatement traités¹⁷, des problèmes apparaîtront avec certitude dans les analyses. Ils conseillent la communauté responsable par les études infométriques et d'indicateurs à continuer à alerter "les producteurs de bases sur l'importance de données plus propres et précises" (p.604).

Critères de qualité

Rittberger et Rittberger (1997) identifient et nomment quatre acteurs dans "la chaîne de l'information" (p.25) : Les **créateurs de connaissance** (les auteurs, brevetés et autres producteurs de connaissance), les **producteurs** (producteurs de bases de données, éditeurs, organisations de propriété intellectuelle), les **distributeurs** (sites, administrateurs de réseaux, grossistes d'information et services) et, par fin, les **utilisateurs**. Tous ces acteurs de la chaîne de l'information doivent faire attention à des normes de qualité et à des nouvelles demandes des clients.

Les normes sont en train d'être créées et développées. Armstrong (1995) indique que dans cette "nouvelle industrie", où le produit primaire est l'information intangible, qualité n'est seulement difficile de définir mais aussi est difficile à quantifier. Cela étant, les auteurs cherchent d'abord se mettre d'accord sur le sujet. Basch (1995, apud. Rittberger et Rittberger, 1997, p.27) rapporte une proposition qui aujourd'hui réunit un grand nombre de personnes, celle du **Southern California Users Group (SCOUG)**. Elle recommande les critères suivants pour évaluer la qualité d'une base de données :

1. Possibilité de déterminer des limites avec respect à géographie, langue et contenu ;

¹⁷ Voir plus détails en "Préparation de données et textes : condition de qualité".

2. Indexation de tous les documents ;
3. Codification de contenus et des types de documents ;
4. Connexion entre des données liées, par exemple, connexion entre actes de conférences et les adresses correspondantes ;
5. Listes précises d'auteurs et de titres ;
6. Non abréviation des titres de revues scientifiques ;
7. Possibilité de recherches complètes en affiliations d'auteurs ;
8. Présence dans la base, au moins, des champs : Auteur, affiliation, titre, source et pays d'origine, date de publication, sommaire et indexation.

Le SCOUG (Basch, 1990, apud Medawar, 1995, p.265) propose aussi dix principes basiques pour la qualité des bases de données :

1. **Accessibilité/facilité d'utilisation** (logiciel de recherche, thesaurus disponible en ligne, permission pour sauver des recherches) ;
2. **Précision** (erreurs typographiques, erreurs d'indexation, intégralité des références) ;
3. **Consistance** (des références dans la base de données, utilisation consistante des champs) ;
4. **Couverture/secteurs d'étendue** (couverture du sujet de la base, étendue, couverture telle quelle est décrite par le producteur) ;
5. **Appui et entraînement à l'utilisateur** (numéro 0800 pour des appels téléphoniques gratuits, personnel entraîné pour attendre des usagers, entraînements gratuits et/ou à des prix raisonnables) ;
6. **Documentation** (impression du thesaurus, manuels actualisés, précis et gratuits et/ou à des prix de coût, aide sensible au contexte) ;
7. **Intégration** (harmonie avec d'autres bases de données du producteur, structure consistante dans les diverses bases de données, dénommer les champs avec consistance) ;
8. **Résultat** (exportation des données selon des formats définis par l'utilisateur, impression sélective de parties du document, exportation pour le courrier électronique) ;
9. **Actualisation** (période d'actualisation, actualité du matériel inclus à chaque fois) ;
10. **Coût/bénéfice** (des formats normalisés et libres et/ou bon marché, caractéristiques du système efficaces en termes de coûts, vitesse, performance et structure de prix).

Dans une enquête réalisée par l'**European Quality in Information Programme (EQUIP)** auprès de 600 usagers de bases de données de différents types, en 12 pays d'Europe (Hollande, Finlande, Royaume Uni, Suède, Danemark, Grèce, Irlande, Slovaquie, Allemagne, Belgique, Suisse et Luxembourg), Wilson (1998) a identifié, dans les dix principes basiques de qualité du SCOUG, les cinq les plus pertinents pour les Européens : **Couverture, précision, accessibilité, actualisation et consistance.**

Dans son travail de revue de la littérature sur la qualité en bases de données Medawar (1995) signale que l'application de méthodes de qualité totale dans des bibliothèques et des services d'information est devenue un thème spécial de la communauté de l'information. Medawar cite l'étude de Fox et al. (1994) qui fait référence à la nécessité d'une meilleure éducation des usagers en stratégies de recherche pour obtenir des meilleurs résultats et une plus grande satisfaction et "décrit les dimensions les plus importantes associées avec la qualité des données : **Précision, couverture, consistance et actualité**" (p.260).

Medawar (1995, p.266) traite aussi de l'initiative du **Centre for Information Quality Management (CIQM)** en Grande-Bretagne. Son objectif est développer une méthodologie pour que les producteurs de bases mesurent la qualité de leurs données et publient les résultats de recherche comme partie de leur stratégie de marchandage institutionnel. Armstrong (1994) rapporte que le CIQM a adopté les principes basics du SCOUG et que la grande partie des réclamations reçues concernait les données eux-mêmes : "Erreurs dans les données, couverture déficiente de revues scientifiques et failles dans les collections disponibles, actualisation déficiente et contrôle déficient de l'indice, représentaient approximativement les deux tiers des messages" (p.46). Les plaintes les plus communes sont : Les problèmes dans la documentation de la base (21%) les déficiences dans les services d'appui aux usagers (10%) et la politique de recouvrements et prix (5%).

Jacsó (1997, p.233) indique que la grande augmentation, en ces dernières années, du nombre de bases électroniques et du nombre de références réunis "peut-être n'a pas été accompagnée par une augmentation similaire de la qualité" (p.237). Ainsi, il recommande l'évaluation systématique du contenu des bases, en termes quantitatifs et qualitatifs comme une des principales activités d'évaluation et de contrôle de la qualité des bases de données.

L'emphase dans le processus

Blakeslee et Rumble Junior (2003) expriment que "le problème est de déterminer combien une base de données est conforme aux critères et principes" (p.36) de qualité. Cela étant, ils affirment que "les processus mis en oeuvre pour assurer la présence de certains critères et principes de qualité dans une base de données" (p.37) sont plus importants que le simple fait de mesurer ou non la qualité de cette base.

Blakeslee et Rumble Junior (2003) décrivent le Standart Reference Data Program (SRDP) du National Institute of Standarts and Technology (NIST). Sa fonction est faire parvenir des

"données référentielles technico-scientifiques évaluées critiquement" (p.35) à des scientifiques et ingénieurs. Ces données sont réunies en 26 bases de différentes disciplines incluant la Biotechnologie, la Chimique, la Physique et la Science des Matériaux. Le NIST observe trois critères : **Précision** - les données doivent être précisément déposées et être identiques aux données finales analysées par les évaluateurs des données ; **Correction** - les données récupérées doivent être les données correctes pour la propriété et la substance sélectionnée ; Et **Fiabilité** - la base de données doit être fiable, ceci veut dire, elle doit toujours donner le même résultat toutes les fois qu'elle est consultée. Un manuel a été produit, **The database process**, pour orienter dans la création des bases de données. Il est divisé en quatre phases : 1) Planification et dessin ; 2) Développement et implémentation ; 3) Révision du produit et édition ; 4) Procédures après-lancement. Les auteurs rapportent que plus de 40 bases de données ont été créées selon ces procédures avec des résultats satisfaisants :

Les participants du projet de base ont une meilleure compréhension de ses rôles et responsabilités. Il y a moins des mal-entendus résultants d'expectatives irréalistes. Des conflits de chronologie sont réduits et les étapes des projets sont accomplies plus proches des délais (Blakeslee et Rumble Junior, 2003, p.44).

La nécessité du processus de planification dans la construction et l'intégration des bases de données est supportée par Baer et al. (1995, p.184) :

S'il y en a une leçon à tirer en cette analyse [le titre de l'article est : "Problèmes dans la création de la base de données OPAC"] est que l'ante vision du processus de planification et aussi l'attention soigneuse aux procédures est essentielle à n'importe quelle activité - mais ils sont spécialement nécessaires pour le processus de production de références qui seront utilisés par un catalogue électronique. Dans un consortium qu'administre une base de données partagée, le processus doit être pris en compte de façon encore plus soignée si l'on ne veut pas être mordu dans les fesses par les dents pointues de l'ignorance et de l'inexpérience (p.184).

Les erreurs dans les bases de données

Erreurs d'indexation

Dans leur révision sur la qualité des bases de données électroniques O'Neill et Vizine-Goetz (1988, p.125) affirment que des contrôles de qualité bien succédés utilisent des méthodologies qui privilégient des techniques manuelles et des techniques automatiques. Entre les techniques manuelles ils citent l'éducation de l'utilisateur sur comment utiliser la base avec efficacité, la

révision des références par un organisme avant leur incorporation à la base et la création d'une voie pour l'indication d'erreurs par les usagers, au moyen du service aux usagers. Entre les techniques automatiques, la conférence par systèmes de données numériques en champs comme ISSN et ISBN, par moyen de formules mathématiques et la création d'algorithmes pour la vérification systématique d'erreurs dans les références, avec correction quand c'est possible. Cela incluent encore des valeurs permises pour le contenu de champs (dates, codes de classification, auteur, titre, mots-clés, séquences de lettres et chiffres) et la consistance entre des différents champs et sous-champs. Des algorithmes existent pour faire la correction automatique ou encore bloquer l'indexation de données inconsistantes ; ils informent l'inconsistance à l'opérateur qui peut la corriger ensuite.

Le 1^o Encontro de aprendizagem do Gedoc [Premier Rencontre d'Apprentissage du Gedoc] (Sd.) indique d'autres types d'erreurs d'indexation très communes : Utiliser singulier et pluriel pour un même concept (quand la recherche est faite sont récupérés seulement les documents au singulier ou ceux au pluriel), utiliser des descripteurs différents pour un même sujet (rend la recherche plus lente et oblige à faire plusieurs recherches sur un même sujet) et "utiliser un trait d'union pour séparer des termes composés quand les termes ne sont pas composés" (génère aussi des résultats partiels dans les recherches).

L'application de méthodes de qualité totale en bases de données électroniques est une tâche que concerne surtout aux producteurs des bases. Medawar (1995, p.261) mentionne article de Mintz (1990) que propose la création de systèmes d'alerte et de messagerie pour que les usagers communiquent des erreurs trouvées en bases de données électroniques et qu'ils puissent être ainsi corrigés.

Dolan (1992) décrit l'importance de l'introduction d'un groupe au département de contrôle de qualité pour l'amélioration du produit base de données électroniques. Casale traite déjà des pratiques adoptées par l'INSPEC qui incluent groupes d'administration de la qualité totale (TQM) pour chaque sujet de couverture de la base, des procédures pour valider les références, correction et vérification d'orthographe dans un dictionnaire qui a débuté avec 70 mille mots mais qui s'accroît à la raison de deux à trois mille mots par mois. L'INSPEC a produit collectivement un manuel d'indexation et encourage des vérifications constantes d'erreurs dans l'indice, le thesaurus, les pages d'accès, le nombre de références et la collection bibliographique.

Erreurs d'orthographe et de dactylographie

Ces types d'erreurs sont analysés en profondeur dans la communauté de l'information. Selon O'Neill et Vizine-Goetz (1988), "à juger par la littérature, les erreurs d'orthographe semblent être les plus communes dans les bases de données" (p.133). Les auteurs se réfèrent à une étude de Bourne (1977) dans 11 bases de données qui a trouvé une grande diversité d'erreurs d'orthographe dans les mots indexés, qui variaient de 0,5% dans la base Biosis à 23% dans la base ABI/Inform.

Damerau (1964), à son tour, a identifié quatre types d'erreurs qui correspondent à 80% des erreurs : **Omission** (le manque d'un caractère, "ausi" au lieu de "aussi"), **insertion** (addition d'un caractère, "ausssi" au lieu de "aussi"), **substitution** (changer un caractère, "aissi" au lieu de "aussi") et **transposition** (inversion de deux caractères adjacents, "masi" au lieu de "mais"). Le 1^o **Encontro de aprendizagem do Gedoc** (Sd.) a indiqué une autre erreur d'orthographe très commune : "L'espace inexistant entre les lettres d'un mot" (au ssi).

O'Neill et Vizine-Goetz (1988) révèlent que beaucoup de ces erreurs plus communes "sont associées avec le clavier de type QWERTY" (p.134). Ils citent Dalessandro (1986) qui a enregistré une augmentation moyenne de précision de 68% en outre de doubler la vitesse de dactylographier quand le clavier commun (QWERTY) a été changé par un clavier de type DVORAK. Pour les erreurs de substitution, par exemple, "la lettre changée est presque toujours dactylographiée avec la même main de la lettre voulue et fréquemment au même niveau horizontal du clavier" (p.134). De la même manière, "presque 84% des erreurs de transposition sont causés par des changements de main droite/main gauche ou main gauche/main droite" (p.134). À enregistrer que les auteurs, d'origine anglo-saxonne, n'ont pas considéré les difficultés supplémentaires posées par les langues latines avec un grand nombre de règles d'accentuation.

Pour Jacsó (1997, p.250) la création d'archives avec des vocabulaires contrôlés comme celui de la base Biosis, avec plus de deux millions de références, dont l'orthographe des mots contrôlés et des noms d'organismes sont vérifiés sont des exemples à suivre. Il cite une étude d'O'Neill et al. (1993) qui indique les champs auteur, date de publication, éditeur et titre comme ces où il y en a le plus d'erreurs. Jacsó (1997, p.252) explique que des erreurs de dactylographie de noms d'auteurs peuvent résulter en que des références ne soient pas trouvées. Cela étant, en des analyses bibliométriques de la productivité d'auteurs, Pao (1989) a découvert que des telles erreurs distordent la distribution à un tel point qu'elle ne se

conformait plus à la loi de Lotka, pendant que la distribution des noms d'auteurs corrigés suivait cette loi. Par conséquent, Jacsó (1993a) recommande qu'avant de faire une recherche l'utilisateur doit parcourir la base pour avoir une idée de ce qu'il trouvera et aussi quelles "techniques de recherche défensive" doivent être utilisées (comme tronquer des mots-clés) pour obtenir des données plus performantes.

Pour éviter les erreurs d'orthographe et de dactylographie plusieurs solutions sont recommandées. L'idéal, affirme Jacsó (1997) est utiliser toujours les archives de vocabulaires contrôlés pour normaliser l'écriture des noms d'auteurs, d'organisations, d'entités géographiques, des descripteurs de thèmes et de sujets, des noms de revues scientifiques, d'année de publication, de langue du document, nom du pays, du département et le code postal, d'entre autres champs. "Ceci est de particulière importance si tous ces champs et éléments sont mis comme points d'accès à la base" (c'est à dire, s'ils sont des champs où des recherches y sont réalisées), conclut Jacsó (p.255). Bourne (apud. Cahn, 1994) considère que les producteurs de bases peuvent aussi détecter avec facilité l'existence d'erreurs d'orthographe avec la création d'indices invertis des champs.

Erreurs d'omission

Medawar (1995, p.262) indique deux articles de Jacsó (1993a et 1993b) sur des erreurs en base de données où en plus des erreurs déjà connues l'auteur aborde aussi des erreurs comme l'indexation de données incorrectes et "l'utilisation de certains champs comme des 'poubelles' pour des informations qui ne s'incluent pas en aucun des autres champs" (p.262).

Jacsó (1993b) affirme que "les pires erreurs sont ces invisibles en ce qui concerne les références" comme, par exemple, l'absence d'informations basiques comme l'année de publication, le type de document, code de langue et code de classification en un grand nombre de références. Il recommande la création par les producteurs de bases de données de systèmes d'alerte et de vérification de toutes les informations considérées basiques pendant et après le catalogage, en donnant attention particulière aux champs qui sont "des points d'accès de recherches" comme, par exemple, langue, pays, département, champ de la connaissance, code postal et mots-clés, parce-que des erreurs en ces items "limitent la récupération".

Mintz (1990, apud. Medawar, 1995, p.263) dénote que le processus de création d'une base de données concerne un grand nombre de personnes. Par conséquent, des erreurs introduites à ce stade incluent en plus des erreurs d'orthographe de tous les types, erreurs d'omission dans les

champs que nécessitent d'être remplis, entrée d'information en champs erronés, mettre des espaces en blanc en locaux erronés et l'utilisation de format erroné dans les champs. En ce cas, des vérificateurs de consistance entre différents champs et sous-champs sont nécessaires pour informer à l'opérateur qu'il y a des champs clés en blanc ou qu'ils sont inconsistants.

Erreurs de duplication

Autre thème exploré en qualité de bases de données électroniques est la détection de doublons. O'Neill et Vizine-Goetz (1988) ont signalé qu'une grande quantité de duplications de références dans une base de données "réduisent sa performance, élèvent les coûts de stockage et maintien et empêchent des recherches efficaces" (p.144). Ainsi, les doublons sont "des questions importantes pour le contrôle de qualité tant pour la part des usagers comme des administrateurs de bases de données" (144). Les doublons ou duplicata sont définies comme "deux ou plusieurs références que représentent le même item bibliographique" (p.144).

Quand divers catalogues sont partagés dans une base unique le problème de duplication devient critique "en fonction des différentes interprétations des normes d'indexation et de la diversité des pratiques de catalogage" (p.144). Cela étant, Sule (1986) recommande que "des bases de données électroniques incluent la vérification de duplicata, en alertant aux opérateurs sur un possible duplication" lors de l'indexation d'un nouveau registre.

Cette tâche de vérification de duplicata est réalisée par des algorithmes pour comparer les références. Ce type de logiciel est très complexe puisqu'ils doivent "compenser pour petites différences des contenus des champs (erreurs d'orthographe, données incomplètes), manque de données et pratiques différentes d'indexation". O'Neill et Vizine-Goetz (1988, p.145) ont révisé une dizaine d'études sur des algorithmes.

Williams et Maclauray (1979) décrivent un algorithme en deux étapes qui réunit des potentiels doublons en comparant titres et dates et, dans une deuxième étape, en comparant auteurs, titres et nombre de pages. Le travail relate que la faille principale de l'algorithme a été en fonction de mots et d'erreurs d'orthographe dans les titres.

Le plus grand problème pour l'algorithme utilisé par l'Université de Californie pour unifier tous ses catalogues bibliographiques, selon Hickey et Rypka (1979) était les variations des champs éditeur, titre et les champs de pagination. Mcpherson et al. (1982) ont expliqué que l'option pour l'automation pour contrôler le critère de la précision s'est donnée en fonction de

"la magnitude du projet. L'intervention manuelle n'était pas pratique" (p.371) et pour cela "des ressources devraient être utilisés en des techniques sophistiquées d'informatique au lieu de procédures éditoriales intensives en travail" (p.371).

Coyle et Gallaher-Brown (1985) rapportent que cet algorithme de l'UCAL opéra en ligne, pendant que le registre est en train d'être digité et détecte, en sa première étape, les références avec les mêmes chiffres d'identification (ISSN, ISBN), date, édition et titre. Ensuite, les références peuvent être unifiées, déclarés non égales et/ou partialement égales et réservés pour traitement postérieur.

Dans l'étape d'unification, Macpherson et al. (1982) signalent que "quand deux références ont des champs identiques, le champ est gardé seulement une fois" (p.375) et est crée une clé de propriété des unités de catalogage. Quand deux ou plus références sont déclarés "partiellement égales" (p.375), des différents algorithmes sont utilisés pour chaque champ - nom d'auteur, nom de l'institution, nom des évènements et noms des titres. Les deux premières étapes ont pour objectif de normaliser les références selon des paramètres prédéfinis. La troisième étape compare les divers éléments en attribuant des poids selon l'importance de l'élément dans le processus d'identification des doublons. Les poids sont additionnés en diverses itérations et à la fin de chacune se décide si le processus continue ou se ferme. Sont analysées 16 possibilités (colonnes) de trois combinaisons (- = inégale, P = partiellement égale et E = égale) en 14 champs de la base (les lignes) comme le montre le tableau 3.

Tableau 3 : Variables pour comparer des références à l'UCAL (Hickey et Rypka, 1979, p.134)

Variable	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Code de reproduction	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Type de registre	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Début du titre	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E	E
Date	E	E	E	E	E	E	E	E	E	E	E		E	E	E	E
Local	-	-	-	E	-	-	-	E	-	-	E	E	-	-	-	-
LCCN	E	E	E	-	P	P	P	P	P	P	E	P	P	P	P	P
ISBN	P	P	P	P	E	E	E	-	P	P	E	E	P	P	P	P
Nombre SuDocs	P	P	P	P	P	P	P	P	E	E	-	-	P	P	P	P
Édition	P	P	-	P	-	-	-	P	P	P	P	P	P	P	P	E
Série	-	-	-	P	-	E	-	P	P	P	P	P	P	P	E	P
Nom d'auteur	-	-	-	P	-	-	-	-	-	-	P	P	E	-	-	-
Page	P	-	-	P	P	-	-	E	E	P	E	E	P	E	P	P
Éditeur	-	P	-	P	-	-	-	P	-	-	-	P	P	P	P	P
Titre haché	-	-	P	E	-	-	P	E	-	P	-	E	P	P	P	E

Hickey et Rypka expliquent que "par exemple, dans la colonne 1 de la figure 3, le tableau de décision, indique que deux références sont classées comme égaux s'ils ont le même code de reproduction, type de registre, data et nombre de code de la Librairie du Congrès des États-Unis et s'ils sont partiellement égales au ISBN, nombre SuDocs, édition et page" (1979, p.134).

Dans la construction du catalogue unifié, Mcpherson et al. (1982, p.372) ont identifié trois champs critiques : a) Le développement d'un format commun pour les références locaux ; b) Le maintien des données quand des nouvelles références sont ajoutés et c) L'unification des références dans une seule base. Macpherson et al. (1982) rapportent les trois décisions plus difficiles prises par le système, sans la supervision humaine :

- (1) Deux références bibliographiques décrivent la même entité bibliographique?
- (2) Lesquelles des références qui composent le registre unifié sont les meilleurs pour être désigné comme le registre de la base et apparaître pour l'utilisateur commun?
- (3) Deux champs (de références qui sont en train d'être unifiés) sont suffisamment similaires qui seulement un d'eux puisse être stocké dans la base de

données? Si les champs sont presque similaires mais pas identiques, lequel doit être préservé? (p.375)

Pour éviter l'élimination ou la non-inclusion de références semblables, Hickey et Rypka (1979), recommandent que le système toujours alerte "quand quelqu'un essaie de créer un registre en duplicité et créa un rapport (log) des références créées même avec cet avis, pour révision manuelle" (p.140).

Hickey et Rypka (1979) entendent que la recherche de cet algorithme a démontré la possibilité "d'identifier des duplicata avec une raisonnable précision et contrôle, en utilisant des clés relativement petites, de 52 bytes" (p.141). Ils ajoutent qu'en étant petite et simple, cette clé "a rendu possible la mise en oeuvre d'une vérification en ligne des doublons, pour aider à prévenir l'addition de nouveaux doublons" (p.141) fait qu'ils classifient comme "une amélioration substantielle" (p.141) dans le processus de qualité de la base.

Déjà Mcpherson de d'autres (1982) créditent une bonne partie du succès au système d'étapes parce qu'il "amplifie les chances d'identifier des références similaires sans augmenter le nombre d'erreurs d'union". Ils déclarent que "l'habileté d'identifier des éléments des références partiellement égales est un facteur clé pour augmenter la précision dans la consolidation de la base" (p.377).

Coyle et Gallaher-Brown (1985, p.78) expliquent que pour être accepté un algorithme doit identifier les différences en détail de chaque registre, des variations dans les pratiques d'indexation et les différences dues à des erreurs. Coyle et Gallaher-Brown (1985, p.79) affirment encore que le processus peut être affiné continuellement par l'examen systématique des références de l'appelé "secteur gris" par des spécialistes qui peuvent déterminer des modifications dans les poids ou dans les étapes de travail de l'algorithme.

Ayres et al. (1984) étudient la viabilité d'utiliser le **Universal standart bibliographic code (USBC)** dans le processus d'élimination de doublons. Ce code a été créé pour identifier un registre à partir d'informations extraites de ses propres champs.

Enfin, Jambu (2000) indique que "la non-qualité d'un système d'informations se caractérise par le taux plus ou moins élevé de données manquantes ou incohérentes" (p.93). Ceci parce que, explique-t-il, un taux anormalement élevé de données manquantes - des données qui devraient être informés mais ne sont pas - produit des calculs et/ou des résultats erronés en chaîne et, à la fin de la chaîne, "le décideur ne sait plus quoi croire" (p.93). La règle d'or dans

l'analyse de données et de textes est qu'exploiter et analyser des ordures donne des ordures comme résultat. En conséquence, l'étape intermédiaire de traitement, préparation, nettoyage et, des uns ajoutent encore, d'enrichissement des données, avant l'analyse proprement dite, est essentielle (Quoniam, 1996).

Les logiciels de bibliothèques

Corte et al. (2002) indiquent trois instruments essentiels dans le processus d'automatisation des bibliothèques :

1. La norme ISO 2709 qui spécifie les réquisitions pour le format d'échange de références bibliographiques entre systèmes informatisés. Elle normalise le structure des références pour rendre possible l'échange d'informations (Corte et d'autres, 2002, pp.37-39).
2. Le protocole Z39.50 a été créé spécialement pour la recherche et récupération d'informations dans les différentes bases de données. Il permet la création de bases de catalogues collectifs virtuels et des recherches pour récupérer des informations et faciliter le processus d'échange bibliographique (Corte et d'autres, 2002, pp.39-43).
3. Le format MARC II, une représentation physique de documents en un moyen lisible par des ordinateurs et capable de contenir des informations bibliographiques de tous les types de matériels. Ses trois composants représentent la structure du document, les indicateurs de contenu et le propre contenu (p.46).

Corte et al. (2002) ont identifié 60 logiciels pour des bibliothèques et 39 logiciels pour des archives. Elles ont fait une enquête avec tous ces fabricants de logiciels. Ont répondu aux questionnaires 24 fabricants de logiciels de bibliothèques et 8 de logiciels d'archives. Le tableau 4 montre les 18 variables analysées et le nombre d'items analysés en chaque variable.

Tableau 4 : Variables pour l'analyse des logiciels de bibliothèques (Corte et d'autres, 2002)

Logiciels de Bibliothèques : Variables analysées	# d'items
Caractéristiques ergonomiques	110
Récupération d'informations	69
Caractéristiques technologiques	44
Traitement technique	44
Circulation et référence	32
Sélection et acquisition	30
Procès de gestion	24
Création de l'entreprise	17
Qualification des ressources humaines	16
Informations générales	10
Caractéristiques générales	8
Services offerts	8
Dissémination de l'information	7
Ambiant en qui le produit opère en les meilleures conditions	7
Périodicité des nouvelles versions du logiciel	6
Méthode de conversion rétrospective adoptée	4
Étendue géographique	4
Ressources humaines	3
Total d'items évalués	443

Note. # d'items = Nombre d'items évalués dans chaque variable analysée.

Entre les huit logiciels d'archives analysés, la recherche de Corte et al. (2002, p.108) a identifié, entre autres points, que les produits "rendraient possible la customisation, l'inclusion ou l'expansion de nouveaux modules en étant disponibles dans la langue portugaise", il y a une préoccupation avec les aspects ergonomiques, "ce qui garantit une navigation conviviale",

l'émission de rapports et statistiques qui "nécessite d'être plus développé". Un point positif est "la possibilité de réaliser des recherches en utilisant des diverses stratégies".

Avec les responsables pour les 24 logiciels de bibliothèque qui ont répondu à l'instrument d'enquête, Corte et al. (2002, pp.154-155) ont identifié entre autres points que la majorité des produits "permet la customisation et l'expansion ou l'inclusion de nouveaux modules" et que les producteurs sont attentifs à tous les items cités dans l'enquête.

Les auteurs concluent que l'ergonomie des logiciels doit devenir le réquisit le plus important, sur lequel doit se baser la décision finale des acheteurs, en raison de sa relation serrée avec la satisfaction des usagers.

Infométrie, scientométrie et des indicateurs de R&D

Spinak (1998) présente la scientométrie comme "un instrument de la sociologie de la science" (p.143) et cite quelques-unes des applications pour ces techniques mathématiques et statistiques :

- Identifier les revues du noyau central de chaque discipline ;
- Estimer la couverture des revues scientifiques secondaires ;
- Adapter des politiques d'exclusion de publications ;
- Identifier les usagers des différentes disciplines ;
- Mesurer l'utilité des services de dissémination sélective d'information ;
- Prédire les tendances de publication de sujets et de thèmes ;
- Élaborer des normes pour la normalisation ;
- Élaborer des processus d'indication, de classification et de confection automatique de résumés ;
- Identifier les tendances et la croissance de la connaissance dans les différentes disciplines;
- Étudier la dispersion et l'obsolescence de la littérature scientifique ;
- Prédire la productivité des éditeurs, des auteurs individuels, des organisations et des pays, d'entre autres (p.143).

Pour Gregolin et al. (2005), la scientométrie, qu'est nommée "la science des sciences" (p.5), vise à développer des

méthodologies pour la construction et l'analyse d'indicateurs, basées sur un apport interdisciplinaire (...) embrasse l'étude des sciences physiques, naturelles et sociales, avec l'objectif de comprendre leur structure, évolution et connexions, de manière à établir des relations des sciences avec le développement technologique, économique et social. Elle s'appuie sur des indicateurs bibliométriques construits à partir de documents publiés en des véhicules spécialisés et concerne des innombrables paramètres, tels que la quantité de publications, les co-auteurs, les citations, la co-occurrence de mots entre d'autres (...) utilisés comment des mesures indirectes de l'activité de recherche scientifique et qui contribuent à la compréhension des objectifs de la recherche, des structures de la communauté scientifique, de son impact social, politique et économique (p.5).

Pereira et al. (1999) caractérisent l'économie de la connaissance par l'utilisation intensive, d'un côté, "de la connaissance originaire de la recherche scientifique : (p.3) et, de l'autre côté, "de la connaissance sur la connaissance, organisée en des bases de données primaires ou d'indicateurs" (p.3). Ils citent que pendant la crise de la vache folle en Europe, le gouvernement français a demandé à l'INIST, éditeur des bases Pascal et Francis, la cartographie de la connaissance existante sur cette maladie et "comme l'INIST possède une base de données bibliographiques avec 12 millions de références, avec contrôle terminologique et méthodes de travail connus comme infométrie, il n'a pas été difficile d'attendre à la sollicitation" (p.3).

Gregolin et al. (2005) constatent que les indicateurs de production scientifique ont été constamment valorisés comme "instruments pour l'analyse de l'activité scientifique et de ses relations avec le développement économique et social" (p.5). Faria (2001) ajoute que "les indicateurs sont une forme de synthèse et d'ajout de la valeur à l'information" (p.9).

Les différents types d'indicateurs

Okubo, 1997 ; Spinak, 1998 ; Narin et d'autres, 1994 ; Courtial, 1990 et Callon et d'autres, 1993 (apud. Gregolin et d'autres, 2005) ont décrit trois types d'indicateurs pour l'analyse de la production scientifique : Indicateurs de production, indicateurs de citation et indicateurs de liaison.

Les indicateurs d'activité/production concernent "le comptage du nombre de publications par type de document (livres, articles, publications scientifiques, rapports), par institution, champ

de la connaissance, pays" (p.7). Ils sont des indicateurs quantitatifs. Ne concernent pas, par exemple, la qualité des publications. Prenant base en ces chiffres, Gregolin et d'autres expliquent que

sont produits des indicateurs de participations au pourcentage, de taux de croissance ou de ratings, des distributions de productivité d'auteurs (loi de Lotka), des distributions de l'utilisation de vocabulaires (distribution de Zipf), des classifications de périodiques, des distributions de revues par sujet (distribution de Bradford) et de demi-vie de publications, d'entre autres (p.7).

Les indicateurs d'impact/citation sont des comptages des citations reçues par un article spécifique. Ils dénotent l'impact d'auteurs ou d'articles, son influence ou visibilité. Gregolin et al. (2005) rapportent que "c'est le moyen le plus connu d'attribuer du crédit à des auteurs" (p.7). De l'indicateur de base, le nombre de citations, sont extraits d'autres comme "le facteur d'impact, l'indice d'immédiateté, l'indice d'impact, l'indice d'activité, l'indice d'affinité et l'indice d'attraction, d'entre autres" (p.7).

Les indicateurs de liaison/relationnels ont à voir avec les cooccurrences d'auteurs, des citations et des mots, des cartes de collaboration, des réseaux de relations entre des auteurs, des institutions et des pays. Sont comptés les chiffres de co-occurrences dans les publications et ainsi mesurés "les efforts de collaboration scientifique" (p.7) et identifiés et cartographiés "la coopération nationale, internationale ou régionale en des différents champs de la connaissance" (p.7).

Leta et Cruz (2003, pp.126-127) utilisent ces trois catégories avec le concept de mensuration : Des mesures d'activité, "le nombre et les caractéristiques des publications" ; Mesures d'impact, "le nombre de fois que les publications sont citées" ; Et les mesures de collaboration ou de coauteurship, "la participation de co auteurs de différents pays dans les publications brésiliennes". Selon Pacheco (2003), les mesures d'activité indiquent "l'intensité de l'activité scientifique" (p.xxvi) à partir du nombre de publications scientifiques. Les mesures d'impact "attribuent importance aux publications par le nombre de fois en qu'elles sont citées" (p.xxvi) et les mesures de coauteurship estiment "l'intensité de collaboration entre les chercheurs des différents pays, par le nombre d'articles publiés avec des coauteurs résidents en ces pays" (p.xxvi).

L'expérience française de l'OST

Pereira et al. (1999) rapportent l'expérience française de l'Observatoire des Sciences et des Techniques (OST). Sa fonction n'est pas de "produire des informations basiques mais d'agréger des couches de valeur à l'information par moyen de son traitement, en créant une nouvelle information" (p.7). L'OST "s'appuya en trois grands piliers : Des bases de données de responsabilité d'autres organismes ; Un ensemble de logiciels pour la construction de bases de données relationnelles et un énorme travail, que demande de l'édition humaine, pour le traitement des données des bases avec lesquelles il opère" (p.7). Ce travail de traitement des données concerne la "reconfiguration et l'enrichissement avec l'introduction d'autres données comme, par exemple, les codes postaux, qui permettent de réunir la production de R&D par des zones géographiques" (p.7). Pour Pereira et al. (1999, p.7) le travail de l'OST doit servir comme inspiration principale pour développer et intégrer des bases de données et indicateurs de la production scientifique brésilienne.

Barré (1997, p.58), directeur de l'OST, décrit un processus en cinq étapes, schématisé par la figure 2, pour la construction d'indicateurs de R&D. La première, l'expression d'une demande (A1) et la volonté manifeste d'utiliser l'élément (indicateur) produit comme une réponse à cette demande (A2). La deuxième étape est la conception (B1), distribution (B2) et production (B3) de l'indicateur. Ensuite vient la construction de la base de données spécifique pour l'indicateur (C) et le travail de calcul des indicateurs primaires à partir des bases de données nationales (D1) et internationales (D2). À ce point, explique Barré (1997, p.60), les bases de données primaires nécessitent d'améliorer l'unification des noms d'institutions. L'étape finale est la recherche socio-économique et méthodologique (E).

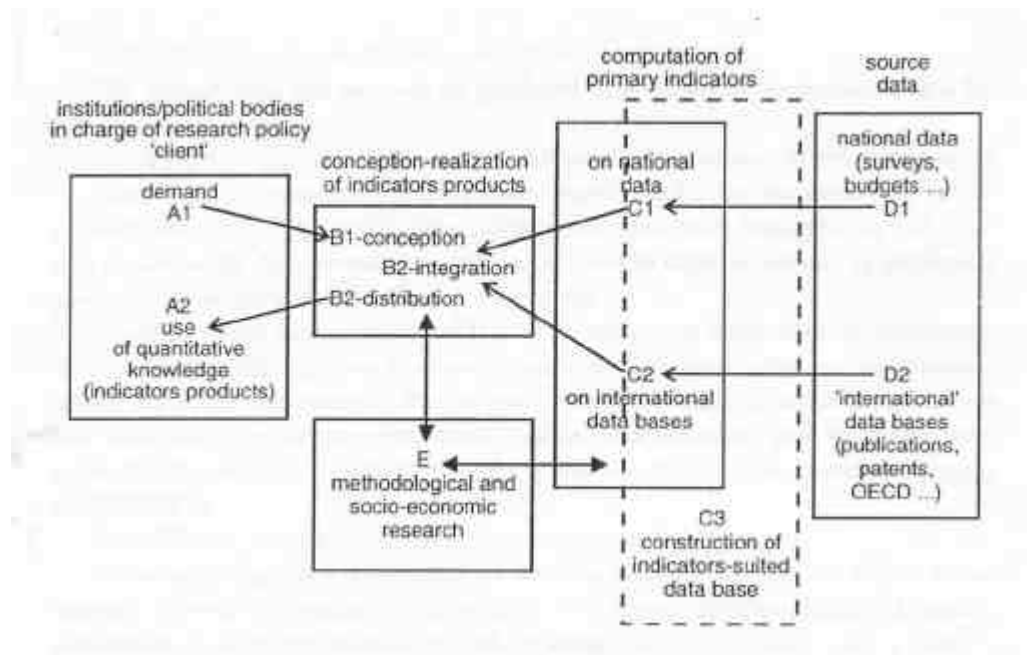


Figure 2 : Le processus de construction d'indicateurs à l' OST (Barré, 1997, p.2)

Barré préconise que la production d'indicateurs doit être le résultat de critères communs d'évaluation déterminés scientifiquement et des modèles conceptuels sur comment des activités humaines comme la recherche et l'innovation se produisent. Il reconnaît que des modèles sont basés sur des hypothèses que nécessitent d'être testées et discutées publiquement.

Barre mentionne que les nouvelles frontières à conquérir sont la production d'indicateurs qui reflètent la diversité des nécessités du processus de décision et la création de moyens pour introduire des nouvelles connaissances et des indicateurs dans ce processus de prise de décision. L'auteur signale la nécessité de rechercher une façon de "mettre en rapport des résultats micro et macro" comme, par exemple, partir d'une vision plus ample et focaliser dans des institutions spécifiques et dans des thèmes détaillés ou, au contraire, d'être capable de partir d'une institution ou d'un thème spécifique et d'arriver au contexte général. Barré mentionne la nécessité "d'accumuler et de capitaliser sur la connaissance déjà générée par les divers efforts de production d'indicateurs et de prospective sur le processus de décision" (1997, p.69). Mais il indique que le plus grand défi est de généraliser les meilleures pratiques de production d'indicateurs de science et technologie pour le plus grand nombre possible de pays.

Indicateurs de R&D au Brésil

L'importance et la valeur des indicateurs reposent sur la possibilité de quantifier des "choses" inintelligibles comme la génération de connaissance. Mais la qualité des indicateurs est liée intimement à la qualité des bases de données.

Pereira et al. (1997) font un "bref diagnostic des bases de données brésiliennes" (p.11) où ils construisent un scénario pas très optimiste. Les indicateurs de production scientifique brésiliens, pour avoir leur origine en bases de données cadastrales, donnent comme résultat des "tableaux et des analyses trop généraux (des sciences exactes, de la terre, des sciences biologiques, des génies, d'entre autres) qui rendent impossible de fonder sur eux quelque décision politique orientée à des thèmes/questions de la connaissance plus spécifiques ou localisées (p.11).

De la même façon, "il n'est pas possible de faire des cartes de connaissance où d'intelligence compétitive à partir de bases de données cadastrales" (p.11) et, "d'un autre côté, il semble hors de propos d'explorer seulement les bases de données internationales pour prendre des décisions stratégiques locales" (p.11). Pour Pereira et al. (1999, p.11) à l'exception des champs des sciences de la vie, de l'énergie nucléaire et de l'agriculture, il y a une couverture déficiente de la production de R&D et de la littérature scientifique brésilienne, par la seule absence d'un contrôle bibliographique national et par la prolifération de bases de données départementales de production scientifique. Cette diversité de bases locales, à son tour, compromet la qualité des données, génère des incompatibilités entre les bases et ne donne pas une visibilité adéquate à la production académique nationale. En conséquence, continuent les auteurs, les analyses générées restent "distantes de la réalité, entraînant des résultats peu constructifs pour améliorer le système de R&D du pays" (p.11).

Gregolin et al. (2005) analysent la production scientifique brésilienne et d'autres pays entre 1998 et 2002 indexée à la base **Science Citation Index Expanded (SCIE)**. Entre les résultats rapportés sont la croissance de la participation brésilienne au nombre total de publications scientifiques, de 1,1% (10.279 articles) en 1998, pour 1,5% (15.846 articles) en 2002, une croissance de 54,2% pendant que la production mondiale s'accrut 8,7%. En termes régionaux, la région Sud-Est continue d'être leader dans la production d'articles indexés avec 77% du total national. Suivent le Sud (15%), Nord-Est (9%), Centre-Ouest (4%) et Nord (2%). Entre les quinze institutions que publient le plus d'articles indexés par la SCIE, treize sont des Universités, entre elles, en ordre décroissante, l'Université de Sao Paulo (USP), l'Université de

Campinas (Unicamp), l'Université Fédérale de Rio de Janeiro (UFRJ), l'Université Paulista (Unesp) et l'Université Fédérale de Minas Gerais (UFMG). Les deux autres institutions qui apparaissent dans cette liste sont la Fiocruz, en neuvième et l'Embrapa, en douzième place.

Gregolin et al. (2005) recommandent "une amélioration permanente des sources d'information nationales, pour viabiliser la construction d'indicateurs bibliométriques plus consistants et appropriés à la réalité du pays" (p.43). Ils citent aussi des initiatives comme "la Plate-forme Lattes, du **Conselho Nacional de Desenvolvimento Científico e Tecnológico**, CNPq [Conseil National de Développement Scientifique et Technologique] et la Base de données de Thèses et Dissertations de l'Instituto Brasileiro de Informação Científica e Tecnológica, IBICT [Institut Brésilien d'Information Scientifique et Technologique], qui méritent d'être mieux explorées pour cette fin" (p.43). Un autre point important indiqué par les auteurs est la nécessité de la création d'un réseau national de recherche, d'élaboration et d'analyse d'indicateurs bibliométriques "pour subsidier le processus de formulation de politiques de science et technologie dans le pays et les départements" (p.43).

Le Ministère de la Science et Technologie (MCT) du Brésil (2005) publie sur l'Internet, sur le site : <<http://www.mct.gov.br/estat/ascavpp/Default.htm>>, les indicateurs suivants de R&D : Généraux, ressources appliquées, ressources humaines, bourses de formation et recherche, d'aide à la recherche, production scientifique, brevets, balance technologique et comparaisons internationales.

À la page de la production scientifique brésilienne, par exemple, on trouve cinq indicateurs : Production scientifique dans le répertoire des groupes de recherche du CNPq, 1998-2001 ; Production technique dans le répertoire des groupes de recherche du CNPq, 1998-2001 ; Orientations conclues et ses respectifs auteurs, 1998-2001 ; Pourcentage d'articles de résidents au Brésil publiés en périodiques scientifiques internationaux indexés par l'**International Science Institute (ISI)**, en relation au total mondial de chaque champ, selon des champs sélectionnés 2000-2002 et articles publiés en périodiques scientifiques internationaux indexés par le ISI et son pourcentage en relation à la publication d'articles par le reste du monde, 1981-2002.

Dans l'item, Comparaisons internationales, le MCT publie des indicateurs de dépenses, de ressources humaines, de dépôt de brevets et de production scientifique. Entre ces derniers il y a aussi des indicateurs plus spécifiques : Nombre d'articles brésiliens, de l'Amérique Latine et

du monde publiés en périodiques scientifiques internationaux indexés par l'ISI, 1981-2002 ; Articles brésiliens, de l'Amérique Latine et du monde publiés en périodiques scientifiques internationaux indexés par l'ISI, selon des champs de la connaissance, 1981-2002 ; Croissance des articles oubliés par divers pays en périodiques scientifiques indexés par l'ISI, 1981-2002 ; Vingt pays avec le plus grand nombre d'articles publiés en périodiques scientifiques indexés par l'ISI, 2002 ; Vingt pays avec la plus grande croissance dans le nombre d'articles publiés en périodiques scientifiques indexés par l'ISI, 1997-2002 ; Articles publiés en périodiques scientifiques indexés par l'ISI, participation en pourcentage en relation au totale mondiale, principaux pays, 2002.

Faria (2001) et Gregolin et al. (2002, p.3) décrivent le processus de production d'indicateurs à partir de la base de données Web of Science (WOS). Celle-ci est une base d'informations bibliographiques et de citations produites par l'ISI qui réunit des articles scientifiques depuis 1945 et s'accroît au rythme de 1,2 millions références par an. Elle totalise aujourd'hui plus de 30 millions de références. Les citations de la WOS sont réunies dans le **Social Science Citation Index (SSCI)** et dans le **Arts & Humanities Citation Index (AHCI)**. Une des caractéristiques de la WOS est qu'elle ne suit pas la tradition des autres bases bibliographiques de considérer seulement l'affiliation du premier auteur des articles scientifiques. Elle cite les affiliations - pays, ville, institution et département - de tous les auteurs bien "qu'il n'est pas possible d'identifier au certès quel auteur appartient à quelle institution".

Cependant Gregolin et al. (2005, p.12) signalent que

même avec ces limitations, autre base bibliographique avec les mêmes attributs pour la production d'indicateurs bibliométriques fiables (multi-disciplinarité, couverture de multi-affiliations et de citations) n'est pas encore disponible et, principalement, avec le même volume d'articles et d'auteurs brésiliens. En 2002 il est possible de compter plus de 15.000 articles indexés par la SCIE contenant au moins un chercheur d'institutions brésiliennes comme auteurs, ce qui renforce son importance pour la construction d'indicateurs de la production scientifique nationale (p.12).

Cela étant, Gregolin et al. (2002) indiquent que "au Brésil, le MCT et la **Fundação de Amparo à Pesquisa do Estado de Sao Paulo** [Fondation de l'Aide à la Recherche de l'Etat de Sao Paulo] (Fapesp) utilisent les données de l'ISI pour en composer ses indicateurs de R&D" (p.4). Avec l'appui de la Fapesp et des Universités Fédéral de Sao Carlos et d'Aix Marseille III, Faria (2001 et Gregolin et d'autres, 2002) a créé une base de données basée en la

WOS, mais beaucoup plus appropriée pour des analyses bibliométriques plus qualifiées, la BiblioWOS.

Dans la base de données BiblioWOS, le registre comme les champs ont été conçus pour faire des analyses bibliométriques automatiques. Les champs ont été maintenus et améliorés. À partir du champ "adresse des auteurs" ont été créés d'autres comme "institution", "département", "ville" et "pays d'affiliation" des auteurs. A partir du champ "citations" ont été créés les champs "auteurs cités", "année de publication" de l'article cité, "source citée", "codes des articles cités" et "nombre d'articles cités" appartenant à la WOS. Dans le champ "auteurs cités" ont été introduits les noms de tous les auteurs des articles cités appartenant à la WOS. Un champ qui classe le registre par la classification de périodiques du ISI a été introduit.

Les avantages de BiblioWOS sur WOS sont évidents : a) Facilité pour récupérer une grande quantité d'information ; b) Organisation des données pour l'analyse bibliométrique automatisée ; c) Récupération de références par les champs de la connaissance par moyen de l'insertion de la classification de périodiques du ISI dans les références ; d) Augmentation de la précision des analyses de citations avec l'introduction des noms de tous les auteurs cités dans les documents indexés par la WOS.

Un des résultats des analyses conduites en utilisant la BiblioWOS a été publié en deux volumes des "**Indicadores de ciência, tecnologia e inovação em São Paulo 2004**" [Indicateurs de Science, Technologie et Innovation à Sao Paulo 2004]. Cette publication a été inspirée, selon Gusmão (2005)

*sur le modèle des principales publications internationales de référence, en particulier les rapports annuels ou bi-annuels d'indicateurs produits par des organismes comme la **National Science Foundation (NSF)**, l'Organisation pour la Coopération et le Développement Économique (OCDE), l'Organisation des Nations Unies pour l'Éducation, la Science et la Culture (UNESCO), la Commission européenne ou encore la **Red Iberoamericana de Indicadores de Ciencia y Tecnologia (RICYT)**.*

Les polémiques sur les indicateurs

Spinak (1998), Velho (2001), Luwell (1999) et Adam (2002) (apud. Gregolin et d'autres, 2002, p.4) expliquent les polémiques sur l'utilisation des indicateurs créés à partir de comptages de publications et citations pour analyser des chercheurs, des institutions et des champs de la science.

Kaplan (1965) a recherché des normes de comportement des citations, en particulier "la possibilité d'avoir des normes pour le nombre de citations par article entre les différentes revues scientifiques et même pour les différents champs" (p.180) de la science. Il a constaté "qu'entre huit à dix pour cent de toutes les citations sont des citations des travaux antérieurs de l'auteur et a peu près 20% de toutes les sources citées dans les revues scientifiques sont des auto citations de la même revue" (p.180).

Kaplan (1965) dit que Garfield lui-même, un des créateurs des indices de citations, alerte sur "le fait évident de que l'importance d'un article ne peut pas être mesurée simplement par le comptage du nombre de citations qu'il reçoit. Garfield fait la distinction entre l'impact et l'importance ou la signifiante" de l'article (p.181).

Velho (2001, p.1) confirme quelques-uns des points soulevés par Kaplan en décrivant une série de problèmes conceptuels et méthodologiques des bases de données bibliographiques et de la base SCIE, en particulier.

- *Il y a une tendance reconnue des chercheurs à s'auto citer (faire de la "propagande" de son propre travail) et à citer leurs amis dont ils connaissent bien les travaux ;*
- *Le système de communication scientifique internationale est très imparfait et l'accès aux articles varie considérablement, ce qui nous amène à nous poser la question de la légitimité de se contenter des citations pour mesurer quelque chose ;*
- *Il y a une barrière linguistique qui très souvent sépare les chercheurs en certaines publications indexées à la SCIE, d'où les citations sont cueillies ;*
- *Des travaux de recherche fondamentale sont beaucoup plus cités que les travaux expérimentaux, ce qui ne signifie pas qu'ils sont plus importants ;*
- *Les pratiques de publication varient entre les champs de la connaissance. On ne pourrait pas dire, par exemple, que la Physique est plus productive que la Mathématique ayant pour base les travaux publiés dans l'une et l'autre discipline. Ce sont des disciplines différentes avec des cultures différentes en ce qui concerne la divulgation dans les revues scientifiques ;*
- *Des articles sont cités sans qu'ils soient lus et analysés soigneusement ;*
- *Les citations sont, dans beaucoup de cas, faites de deuxième main, c'est à dire, ont été prises dans un autre article, sans avoir été lues et sans donner crédit au travail qui, en premier lieu, les a cités ;*

- *Les citations sont, dans beaucoup de cas, insérées après que le travail soit fini, presque comme une 'décoration' pour suivre le 'rituel' académique (p1).*

Velho (2001, pp.2-3) parle encore d'autres problèmes. Les classifications scientifiques opèrent des comptages de production scientifique avec des "vices d'origine, [comme] la priorité au Premier Monde et des présuppositions simplistes concernant l'Amérique Latine" (p.2). La base SCIE, par exemple, privilégie des revues scientifiques publiées par les pays de la "source principale", soit, États Unis, Japon, Allemagne, Angleterre et France. Seules 15 revues scientifiques brésiliennes figurent parmi les 5.891 revues indexées. Le résultat de cette sous-représentation régionale est la sous-évaluation de la science dans le reste du monde, comme par exemple, en Amérique Latine.

Entre les vices d'origine, explique Velho (2001), très souvent les prémisses sur la production qualitative ne sont pas vraies. On ignore que "la publication formelle est seulement l'un des types de communication en science" (p.2), différent, par exemple, de l'apprentissage, de la connaissance tacite (Nonaka et Takeuchi, 1997), du contact journalier des équipes. De la même façon, tous les articles n'ont pas "la même contribution à la connaissance" (p.2) et entre les chercheurs, les comportements différents selon les champs de la connaissance, les pays et leurs activités, s'ils militent dans la recherche basique ou appliquée. Pour Velho (2001) "la production de nouvelles connaissances scientifiques comme celles [...] publiées dans les revues" (p.2) ne sont pas le seul objectif de la science. Il y en a des autres comme :

La contribution pour la solution des problèmes nationaux, régionaux, locaux ; La contribution pour le développement technologique ; La formation et l'entraînement de nouveaux chercheurs et du personnel qualifié pour les professions ; L'augmentation du prestige national et international d'un pays (ce qui, dans les sociétés modernes, a une connotation de pouvoir) ; D'être le garant de l'autonomie d'un pays dans un secteur déterminé et d'influencer sur la vision du monde et de la société en général (p.2).

Gregolin et al. (2002, pp.4-5) enregistrent encore d'autres problèmes qui rendent difficile la précision des indicateurs créés à partir des citations de la WOS. Il y a des limitations de :

1. Contenu. Seulement les premiers auteurs sont cités et il n'y a pas de classification par des champs de la connaissance. Les informations sur tous les autres chercheurs sont perdues ;
2. Structure inadéquate pour le traitement automatique. Des auteurs, des citations, des entités, des pays (dans le cas des citations, par exemple ; Les informations du chercheur, périodique, année, volume et page, dans un seul champ. Dans le cas des affiliations, les données de l'entité d'origine, département, état, ville, pays et code postale aussi dans un seul champ) ;

3. Accès. Limitation du nombre de références recouvrées et sélection très longue des références pour téléchargement.

Dou (1989) indique qu'il y a des méthodes nombreuses pour la création d'indicateurs, "mais ce qu'on ne doit pas faire est de créer des indicateurs avec le seul objectif de les produire" (p.28). L'auteur argumente qu'en observant les sciences et les techniques nationales ou régionales, le plus important est de définir des indicateurs généraux et globaux et d'utiliser des indicateurs plus précis pour appuyer les activités d'intelligence.

Pour Jakobiak (1991), il y a "un certain abus" (p.148) quand on réduit l'intelligence compétitive (ou économique) à un problème d'indicateurs. Ceci est "extrêmement périlleux et peut décrédibiliser la nouvelle discipline" (p.148). Les indicateurs, continue Jakobiak, ont le pouvoir de "rendre explicites des faits en rien évidents a priori" (p.148) et de mesurer des effets en utilisant pour cela des extraits de bases de données, de déterminer les principaux thèmes de recherche et, en suivant, d'analyser les paires présentes dans les références de la base pour découvrir les réseaux relationnels entre les différents domaines du sujet étudié et encore suivre son évolution pendant le temps. Les techniques scientométriques proportionnent une prospective de moyen et long terme pour définir des priorités et des études des secteurs mais, conclut Jakobiak (1991, p.149), elles doivent être utilisées toujours d'une manière ponctuelle.

LA CRÉATION DE SYSTÈMES
D'INTELLIGENCE À L'EMBRAPA

La création de systèmes d'intelligence à l'Embrapa

L'Entreprise Brésilienne de Recherche Agronomique (Embrapa), attachée au **Ministério da Agricultura, Pecuária e Abastecimento** [Ministère de l'Agriculture de la Production Animale et de l'Approvisionnement], a été créée en 26 avril 1973. Sa mission est de viabiliser des solutions pour le développement durable de l'espace agricole brésilien, focalisés sur l'agronégoce, par moyen de la génération, l'adaptation et le transfert de connaissances et technologies, dans l'intérêt de plusieurs segments de la société brésilienne.

L'Embrapa agit par l'intermédiaire de 37 centres de recherche, trois services et 11 unités centrales, étant présentes dans presque tous les états de la fédération, dans les plus différentes conditions écologiques. Pour devenir l'une des plus grandes institutions de recherche du monde tropical, l'entreprise a investi beaucoup dans la formation des ressources humaines. Elle possédait en 2005, 8.619 employés, desquels 2.221 étaient chercheurs, 45% avec Diplôme d'études approfondies et 53% avec Doctorat. Son budget est de l'ordre de R\$ 877 millions annuels (E\$ 400 millions). L'Embrapa est chargée de la coordination du Système National de Recherche Agronomique (SNPA), constitué par des institutions publiques fédérales et départementales (états), des universités, des sociétés privées et des fondations qui collaborent dans l'exécution de recherches dans les différentes régions géographiques et champs de la connaissance scientifique.

L'Embrapa adopte le système Embrapa de gestion (SEG) structuré dans six macro-programmes : 1. Grands défis nationaux ; 2. Compétitivité et développement durable sectoriel ; 3. Développement technologique incrémentiel ; 4. Transfert de technologie et communication ; 5. Développement institutionnel et 6. Aide au développement de l'agriculture familiale et au développement durable du milieu rural.

L'agronégoce et le futur

Le Produit Intérieur Brut du Brésil se répartit de la façon suivante : 10% du total se rapporte à l'Agriculture, 40% à l'Industrie et 50% aux Services. Néanmoins, Gasques et al. (2003) précisent que lorsqu'on intègre toutes les affaires agricoles : Les facteurs de production agricole, l'agriculture et la production animale et forestière en soi-même, l'industrie de

transformation de ces produits et leur distribution, l'agronégoce arrive à une production équivalente à 33,8% du PIB brésilien (p.9).

Outre le fait de répondre pour environ un tiers au PIB du Brésil, l'agronégoce brésilien a été "le secteur de l'économie qui a contribué le plus pour la formation du surplus de la balance commerciale du pays" (p.11). "En 2002, il a répondu par 41,1% des exportations" (p.11). La gamme des produits s'accrut. Elle a passé du traditionnel café, tabac, papier et cellulose, sucre, soja, jus d'orange et des poulets pour incorporer d'autres produits comme les fruits tropicaux et tempérés (mangue, papaye, melon, raisin et pomme), le coton, les viandes porcine et bovine. La qualité des produits a élargi les marchés traditionnels et a conquis de nouveaux comme la Chine, la Russie, les pays du Moyen-Orient, le Chili et l'Indonésie. Des excédents commerciaux de l'agronégoce brésilien ont grandi de dix milliards de dollars par an en 1990 pour 34 milliards de dollars par an en 2004, selon le Ministère de l'Agriculture de la Production Animale et de l'Approvisionnement (2005).

L'agronégoce brésilien est aussi le plus grand employeur du pays. Gasques et al. (2003, p.13) rapportent que l'Institut Brésilien de Géographie et Statistique (IBGE) calcule en 17,4 millions, les emplois directs dans le secteur agricole, l'emploi rural, ce qui représente 24,2% de la population économiquement active (PEA) brésilienne. "Pour mieux dimensionner cette participation, il faut garder en mémoire que la construction civile, considérée grande absorbante de main d'œuvre, occupe 7% de PEA" (p.13). Si nous incluons aussi les emplois de toute la chaîne de négoce tournés vers l'agricole, le total d'emplois générés arrive à 27 millions. L'agronégoce représente pourtant un tiers du PIB et plus d'un tiers des emplois brésiliens, 37% de PEA (Embrapa, 2004, p.10).

Un autre fait à noter est la composition de la croissance de la production agricole brésilienne. Entre 1990 et 2003, la production de grains a passé de 57,8 millions de tonnes à 120 millions de tonnes, une croissance de 108% dans le volume. Quand nous analysons la surface plantée, de 1999 à 2003 elle s'est agrandie de 37,8 millions d'hectares pour 47 millions d'hectares, une croissance de 24%. Gasques et al. (2003, p.28) soulignent que la production animale se développe aussi "à des taux élevés" (p.28). Entre 1990 et 2002, la production de volaille a augmenté de 223%, de bœuf (viande) de 125%, de poisson de 68% et de lait de 45%. Tous ces chiffres indiquent que la base principale de la croissance de la production agricole brésilienne n'a pas été l'accroissement de la surface plantée en des nouvelles régions comme, par exemple, en Amazonie, mais l'augmentation de la productivité.

Bonelli (2002) mentionne que l'impact de la recherche agronomique est celui d'une "révolution invisible" qui apporte "l'inclusion sociale" surtout dans les petites et moyennes villes de l'intérieur du pays. Gasques et al. (2003, p.12) soulignent que "la grande majorité de petites villes brésiliennes à son économie fondée sur l'agronégoce". Bonelli (2002, p.36) a identifié "une forte association entre le niveau du revenu agricole et l'indice des conditions de vie (ICV)". Et que le développement économique et social associé à l'agricole a des conséquences positives sur la "génération de revenus, la croissance de la population, la collecte d'impôts et sur le développement humain dans des espaces geo-économiques sélectionnés" (Bonelli, 2002, p.33). En étudiant des régions de frontière agricole et d'expansion récente comme Balsas, au Maranhão, le Pôle Açu-Mossoró au Rio Grande du Nord, Petrolina au Pernambuco, Rondonópolis au Mato Grosso et Rio Verde au Goiás, Bonelli conclut que

Balsas a grandi à un taux annuel moyen de 8,49% entre 1975 et 1996 et l'état du Maranhão a grandi seulement 2,01%; Au sud du Piauí, où se développent les plantations du soja, le taux de croissance moyen annuel a été de 8,28% et le taux de croissance moyen de tout l'état a été de 3,64%; Petrolina, région de production de fruits par irrigation, a grandi à un taux annuel de 13,27%, tant que le taux moyen de croissance de l'état de Pernambuco a été de 3,73% (Bonelli, de 2002, apud. Gasques et d'autres, 2003, p.28).

Gasques et al. (2003, p.23) soulignent aussi l'importance de l'agriculture pour le succès du programme brésilien de stabilisation économique de 1994/1995, le **Plano Real**. L'agronégoce a eu un effet positif qui a été connu comme "l'ancre verte", la stabilisation des prix des aliments au niveau du consommateur.

Barros et Rizzieri (2001) ont étudié l'évolution des prix des principaux types d'aliments dans la période 1975-2000. Ils révèlent qu'il y a eu une chute généralisée des prix de tous les aliments de 5,25% par an (p.11), ce qui a représenté "un taux moyen de croissance de 7,56% par an, du pouvoir d'achat du salaire moyen d'un maçon salarié dans la ville de Sao Paulo" (p.15) ce qui signifie, indiquent les auteurs, *l'un des plus grands transferts effectifs de revenu envers les couches les plus humbles de la population brésilienne.*

Gasques et al. (2003, p.42) indiquent encore que "la technologie a eu un rôle décisif" dans le succès de l'agronégoce et "l'Embrapa a eu un rôle de leader dans la génération et la diffusion d'innovations dans le secteur agricole". L'entreprise a eu une participation capitale dans "l'expansion de la frontière agricole pour les régions des cerrados (savanes brésiliennes)" avec

"le développement pionnier de variétés de soja", ce qui a introduit "une rupture technologique clé" (p.42) et a permis de faire de l'agriculture dans des "régions avant considérées impropres à l'agriculture. Cette innovation signifie que le Brésil est aujourd'hui l'un des rares pays du monde avec des possibilités effectives d'élargissement de sa surface plantée" (p.31).

Comme le signale Sílvio Crestana, directeur-président de l'Embrapa (Embrapa, 2005a, p.7),

le Brésil peut encore multiplier par trois sa production de grains, passant de 120 millions à environ 350 millions de tonnes, ceci parce que presque 90 millions d'hectares de terres cultivables n'ont pas été utilisés encore. La recherche indique aussi que ce volume pourra être encore plus grand, puisque 30% des 220 millions d'hectares aujourd'hui occupés par des pâturages peuvent être incorporés à la production agricole (p.7).

Pour Penteadó et al. (2002),

l'Embrapa mise sur un avenir prometteur pour l'agronégoce brésilien. Sa compétitivité se définira, chaque fois plus, par sa capacité d'incorporer, de façon continue et soutenue, des innovations technologiques qui permettent de suppléer aux exigences croissantes du marché interne et de défier les subventions des autres concurrents et la tendance historique à des prix décroissants dans le marché international de produits agricoles.

Ainsi, l'Embrapa met en oeuvre une procédure systématique d'analyse et d'ajustement de son Programme (Agenda) Prioritaire de R&D, de façon à se concentrer davantage sur le développement d'innovations que viabilisent la compétitivité et le développement durable du secteur agricole et des négoce agricoles de petite portée, en contribuant à l'augmentation du revenu, de l'emploi et de l'insertion sociale, à la réduction de risques et à l'augmentation de la précision, à l'accélération des processus d'information et de décision et à l'agrégation de la qualité et de la valeur aux produits de l'agronégoce brésilien.

Les objectifs stratégiques de l'Embrapa

Les objectifs stratégiques de l'entreprise selon son IV Plan Directeur (2004) sont les suivants :

Consolider les bases scientifiques et technologiques, promouvoir des innovations et des arrangements institutionnels appropriés pour développer la compétitivité et la durabilité de l'agronégoce, en bénéfice de la société brésilienne ;

Élargir et fortifier les bases scientifiques, promouvoir l'innovation technologique et les arrangements institutionnels appropriés pour développer les capacités productives des petits producteurs et les entrepreneurs, avec un développement durable et compétitif ;

Fortifier les bases scientifiques, promouvoir l'innovation technologique et les arrangements institutionnels appropriés pour rendre propice la sécurité alimentaire, la nutrition et la santé de la population ;

Développer et fortifier les bases scientifiques et promouvoir l'innovation technologique et les arrangements institutionnels appropriés pour rendre propice l'utilisation durable des biomes ;

Promouvoir l'avance de la frontière de la connaissance scientifique et technologique dans des sujets stratégiques pour l'Embrapa.

La politique de communication de l'Embrapa

L'Embrapa a adopté en 1996 une politique de communication d'entreprise, révisée et élargie en 2002 qui, affirme Silva (2004. p.1), "a inséré stratégiquement la communication dans sa politique d'administration". La politique adopte un concept spécifique pour la communication d'entreprise (Embrapa, 2002, pp.28-29) qui

consiste en un processus de gestion qui intègre toutes les activités orientées pour les relations entre une organisation et son environnement interne et externe. Sa responsabilité fondamentale est de créer et maintenir des flux d'informations et d'influence réciproque entre l'entreprise, ses publics d'intérêt et la société en général.

*Comme **processus** d'administration, la communication obéit à une dynamique propre et est organiquement associée aux procédures globales de prise de décisions. Elle se définit par l'intégration, donc une articulation productive, de toutes les activités focalisées sur les relations avec les publics d'intérêt, de manière à ce que soient garanties des conduites uniformes pour toute l'organisation, en évitant l'éventuelle fragmentation de l'image de l'entreprise. En étant **permanente et systématique**, la communication d'entreprise se place comme une ressource stratégique du processus de planification de l'entreprise, en agissant donc, en synergie avec les actions institutionnelles et de **marketing**. Dans la recherche pour satisfaire des **demandes, des nécessités et des attentes** des publics d'intérêt de l'entreprise, elle est en consonance avec la Politique Globale d'Administration et ses principes.*

Ce concept rejette la traditionnelle vision de la communication d'entreprise comme une simple annexe du processus de gestion et la définit, stratégiquement, comme un instrument d'intelligence de l'entreprise (pp.28-29).

Et continue:

Comme instrument d'intelligence compétitive, la communication d'entreprise exige planification, qualification et adoption de pratiques modernes de collecte,

traitement, récupération et transmission des informations. Surtout, elle doit être intégrée au processus de prise de décision et ne pas venir à la traîne, comme dans le modèle traditionnel où elle fonctionnait comme simple annexe du processus de gestion. Dans une institution de recherche, admettre la communication d'entreprise comme stratégique, implique de disposer de cette compétence dans toutes les instances de l'organisation, soit au moment de prospection des demandes ou de la construction des scénarios, soit dans l'interaction avec les publics d'intérêt qui participent des projets et des solutions acheminées par l'entreprise ou dans la recherche de l'interface nécessaire avec la société (pp.14-15).

Et encore :

La communication de l'Embrapa a évolué vers un système d'intelligence d'entreprise soutenu dans la planification, dans la veille systématique et continue, soit pour les activités de communication qui se développent, soit pour l'insertion de l'entreprise dans son environnement externe. De cette façon elle a cherché à mettre en oeuvre des banques de données intelligentes qui incluent les attributs principaux de ses publics d'intérêt (son profil, ses attentes, ses demandes, entre autres). De la même manière, elle a utilisé des méthodologies pour évaluer les résultats de ses actions et recourir à des recherches et à des audits (d'image, de présence dans les médias, des niveaux de satisfaction des clients, entre autres), cherchant à définir des indicateurs de performance et même d'anticiper des tendances pour accomplir de nouveaux défis (p.27).

Et, finalement :

La définition de ce nouveau profil pour la communication d'entreprise à l'Embrapa nous renvoie, obligatoirement, vers la construction et la mise en oeuvre d'un modèle qui l'envisage, prioritairement, comme instrument d'intelligence économique et de l'entreprise. Cela signifie un travail de construction de relations avec les publics d'intérêt à partir d'une perspective stratégique, c'est-à-dire, attachée au processus de prise de décisions et soutenue par des actions et des méthodologies syntonisées avec cet objectif.

Dans ce sens, l'adoption d'une culture tournée vers la veille permanente et systématique des activités de communication, du macro-ambient et du profil des publics d'intérêt, devient fondamentale. Elle sera soutenue par le développement de méthodologies qui permettent la construction de scénarios qui évaluent la relation coût x bénéfice, sous le point de vue institutionnel et commercial de chaque action, produit ou stratégie de communication développée par l'entreprise et auditent l'efficacité de l'insertion de l'Embrapa dans le marché et dans la société.

La pratique de la recherche, essence du travail de l'entreprise, doit s'élargir aux activités de communication d'entreprise, en vue de la conviction qu'il est important de produire des solutions et de la technologie, mais aussi d'analyser leurs impacts et les niveaux de satisfaction des publics auxquels elles se destinent.

Encore plus : Il est nécessaire de développer des projets en visant à découvrir des nouveaux formats, des nouveaux canaux et véhicules, ainsi que d'enquêter pour voir si les discours et le langage s'ajustent à l'héritage socioculturel, économique et éducationnel de l'audience.

La communication d'entreprise comme un instrument d'intelligence d'entreprise exige, encore, le modelage, l'implantation et le maintien de banques de données intelligentes qui prennent compte des attributs des publics d'intérêt, des tendances du marché, de la société et de l'industrie de la communication. Il doit inclure aussi l'organisation d'un réseau interne, soutenue dans les nouvelles technologies qui permettent le partage et la circulation rapide de données, d'informations et de connaissances, visant à les rendre rapidement disponibles au fur et à mesure que seront nécessaires pour subventionner des actions d'intervention dans le marché et dans la société, qui légitiment et garantissent la durabilité institutionnelle de l'Embrapa (pp.83-84).

Intelligence compétitive à l'Embrapa

Avec le support de la politique de communication, en janvier 2001, l'Embrapa a initié un projet stratégique pendant trois ans, avec fin en décembre 2003, conduit par le doctorant, avec le titre ; "Intelligence Compétitive et Gestion de la Connaissance à l'Embrapa". L'initiative reconnaissait formellement que l'organisation

est mûre et commence l'application de techniques de collecte et d'analyse systématique des informations. Avec ce projet, l'entreprise élargit et organise de façon systématique ses investissements dans ce secteur stratégique, intégrant l'intelligence compétitive (IC) et la gestion de la connaissance (GC) aux processus de planification et de gestion stratégique (Penteado, 2000).

Le projet avait comme objectif de créer des réseaux d'intelligence compétitive et de gestion de la connaissance à Embrapa,

adapter la méthodologie existante de création de systèmes d'intelligence compétitive pour la réalité de l'Embrapa, créer des indices et des indicateurs d'évaluation pour les objectifs stratégiques (OEs) travaillés [par le Bureau de Communication Sociale - ACS], générer des méthodologies pour création de réseaux d'IC pour les autres OEs et soutenir le développement de systèmes d'informations et d'intelligence compétitive dans d'autres secteurs comme de l'intelligence scientifique et technologique, de l'intelligence de marchandage, du produit, des concurrents et du macro-ambient. C'est-à-dire, contribuer au succès des objectifs stratégiques du Modèle de Gestion Stratégique (MGE) et ouvrir, en même temps, le chemin pour l'utilisation de l'intelligence compétitive dans l'entreprise (Penteado, 2000, p.8).

Et justifie Penteado (2000) :

Tyson (1998) établit quatre phases dans l'évolution de l'intelligence compétitive dans les sociétés, vis-à-vis la planification stratégique. Dans la première phase les entreprises ont peu ou aucune planification stratégique. Ces sociétés n'y auront pas d'intentions ou des activités d'intelligence. Dans la deuxième phase les entreprises ont la planification stratégique comme une activité isolée. Dans la troisième phase les entreprises ont un processus de planification stratégique bien développé. En ce moment, affirme Tyson, l'efficacité de la planification et de la mise en oeuvre des plans exigera l'intelligence et les processus commencent à s'implanter. Dans la quatrième phase, la plus avancée, une interface systématique est créée entre les processus de planification stratégique et d'intelligence. Dans cette étape, l'intelligence devient complètement intégrée à la vie de l'entreprise et la planification stratégique est efficace et continue.

L'Embrapa dispose d'un système de planification stratégique forte et opérante. Examiné à la lumière de l'histoire de l'IC dans les organisations, il se place à la troisième phase décrite par Tyson (1998), où le développement de la stratégie d'entreprise mène naturellement à la nécessité d'une veille stratégique de son environnement, de ses concurrents et de ses affaires. L'entreprise a adopté, ces dernières années, des mesures qui visent l'intégration des systèmes de planification stratégique, en créant des initiatives d'intelligence.

Certaines d'entre elles peuvent être citées comme l'établissement, en 1998, d'un Laboratoire Virtuel à l'Extérieur - États-Unis (LABEX) ; Le III Plan Directeur de l'Embrapa (PDE) 1999-2003 (Embrapa 1998), qui a placé la veille technologique (p. 34) comme un des projets nouveaux de l'entreprise et, plus récemment, le modèle de gestion stratégique (MGE) qui a intégré des mécanismes de gestion ; Le système Embrapa de planification (SEP), le système d'évaluation d'unités (SAU), le plan annuel de travail (PAT) et le système de planification, d'accompagnement et d'évaluation de résultats du travail individuel (SAAD). Le MGE a défini aussi un objectif stratégique, de n.º 17, "Mettre en oeuvre la gestion de l'information pour soutenir les activités techniques et scientifiques, administratives et de gestion".

Grâce à sa forte planification stratégique et au MGE, l'Embrapa est extrêmement bien placée pour initier le développement de réseaux d'IC pour la gestion et en d'autres secteurs, comme l'intelligence technologique, tournées vers les négoce et le macro-ambient. Les réseaux d'IC seront comme des extensions naturelles des procédures d'administration stratégique, s'intégrant aux efforts d'amélioration organisationnelle, perfectionnant la qualité des décisions, permettant une plus grande rapidité de réponse et une moindre duplication d'efforts. La stratégie d'implantation sera "d'apprendre en faisant", articulée avec des coordinateurs administratifs, leaders de projets et sous-projets, gérants d'objectifs stratégiques, dirigeants du MGE, avec les directeurs des unités centrales et décentralisées et avec la direction-exécutive, publics prioritaires du projet. Les systèmes d'informations stratégiques résultants de cette initiative contribueront pour que l'entreprise améliore sa performance, dans des termes de qualité totale et aussi en concernant les critères d'évaluation de la gestion publique qui intègrent le programme de qualité et participation dans l'administration publique (QPAP). Ils

permettront à Embrapa d'élever substantiellement sa performance et son concept dans des évaluations de ce type (p.1).

Le projet opérait avec deux sous-projets : 1. Création d'un système pilote d'intelligence compétitive et de gestion de la connaissance pour la communication de l'Embrapa, conduit par le doctorant et 2. Développement d'un modèle pour cartographier les technologies produites par l'Embrapa en Amazonie : Enquête des informations et applications bibliométriques conduit par la bibliothécaire-chef de l'Embrapa Amazonie Orientale (CPATU) Rubenise Farias Gato.

Parmi les résultats du sous-projet 1. Création d'un système pilote d'intelligence compétitive et de gestion de la connaissance pour la communication de l'Embrapa, nous citons :

- Le mémoire de DEA du doctorant à l'Université du Sud Toulon-Var : "Systèmes d'Intelligence Compétitive pour la Gestion Stratégique dans une Institution de Recherche et Développement" ;
- Révision et mise à jour de la base des chercheurs et compétences de l'Embrapa - le Guide de Sources - ainsi que la génération de produits d'intelligence compétitive traités en détail au chapitre "Analyse stratégique des compétences de l'Embrapa" ;
- Révision et mise à jour de la base de clippings de l'Embrapa avec la création d'une nouvelle version du Clipping Électronique de l'Embrapa en Lotus Notes qui permet la réalisation de recherches et le développement d'analyses de la performance de l'effort de communication sociale de l'entreprise traités au chapitre "Analyse de la performance de l'Embrapa dans les médias" ;
- Développement et mise en oeuvre d'un système de recherche d'opinion par courrier électronique pour réaliser des recherches et évaluer des actions et la satisfaction entre les divers publics de l'Embrapa.

D'entre les résultats du sous-projet 2. Développement d'un modèle pour cartographier les technologies produites par l'Embrapa en Amazonie : Enquête des informations et applications bibliométriques nous citons :

- Le document d'orientation pour la création du "Noyau Thématique d'Informations pour l'Agronégoce" qui utilise des stratégies d'intelligence compétitive pour la "collecte, gestion, analyse et dissémination d'informations sur les environnements compétitifs,

concurrentiels et organisationnels, visant à soutenir le processus de décision et à atteindre les objectifs stratégiques du centre". Le noyau considère deux fonctions du secteur d'information : Une rapportée aux collections et quantités (bibliothèque, mémoire institutionnelle, conservation et stockage) et autre liée au flux d'informations (dissémination et transfert d'informations) toutes les deux exigeant qualification professionnelle spécifique, reconnaissance de leur importance et complémentarité. Le noyau prête des services pour R&D et pour l'Agronégoce et s'occupe aussi du développement de méthodologies et de techniques propres du secteur d'information ;

- Évaluation de la collection de périodiques de la bibliothèque dans la période 1998-1999, sur un total de 618 titres ;
- Révision et mise à jour de la base **Collection documentaire** de l'Embrapa Amazonie Orientale et d'autres unités de recherche en Amazonie avec l'identification et la correction de doublons et d'erreurs d'indexation, d'orthographe, de dactylographie et d'omission. Entre novembre 2000 et novembre 2003, le nombre de références propres et révisées, contenues dans la Collection documentaire de l'Embrapa Amazonie Orientale, a augmenté de 45%, de 20.863 à 30.388 notices.

Le projet a été encore responsable de la réalisation de cinq séminaires où ont été présentés des travaux réalisés par les techniciens et directeurs de recherche associés et responsables du projet :

1. Le "I Séminaire International Outils pour l'Intelligence Compétitive", du 29 septembre au 3 octobre 2003, à Brasília, Brésil. Parmi les conférenciers on peut citer le professeur Henri Dou, de l'Université d'Aix Marseille III, sur "Des Systèmes de **mind mapping** et de détermination de la vision d'une organisation et le logiciel Matheo Analyser" ; Sri Manullang, doctorante de l'Université d'Aix Marseille III, sur "Des indicateurs scientifiques pour l'appui à des programmes de développement et le logiciel Matheo Patent" ; Roberto Penteado, doctorant de l'Université du Sud, Toulon-Var, sur "De la création de bases de données au développement de systèmes d'intelligence pour l'organisation" ;
2. Le Séminaire "Gestion des Informations et de la Connaissance dans l'Embrapa Amazonie Orientale", du 11 au 14 novembre 2002, à Belém, Pará, Brésil. D'entre autres, ont été conférenciers, le professeur Luc Quoniam, de l'Université du Sud, Toulon-Var, sur "L'intelligence compétitive et la gestion de la connaissance et son importance pour les organisations" ; Le professeur Lena Vânia Pinheiro Ribeiro, de l'Université Fédérale Fluminense et de l'Institut Brésilien d'Information en Science et Technologie (IBICT), sur "L'expérience de construction de la bibliothèque virtuelle agricole en Amazonie" ; Roberto Penteado, doctorant de l'Université du Sud, Toulon-Var, sur "L'application de la bibliométrie dans l'analyse stratégique des compétences de l'Embrapa Amazonie

Orientale. 2002" ; Rubenise Farias Gato, leader du sous-projet 2, sur "Les premiers résultats du sous-projet "Développement d'un modèle pour cartographier les technologies produites par l'Embrapa en Amazonie";

3. Le "Séminaire SEA Gestion Stratégique des Compétences de R&D à Embrapa", le 9 octobre 2002, à Brasília, Brésil. Ont été présentés par le conférencier, Roberto Pentead, doctorant de l'Université du Sud, Toulon-Var, les résultats du travail du sous-projet 1, développés conjointement avec le Département de Recherche et Développement (DPD), l'analyse des compétences de l'**Agenda** de R&D de l'Embrapa avec 2.355 chercheurs ; 2.173 actifs et 182 inactifs depuis 1997, avec l'objectif de favoriser la gestion stratégique des compétences de R&D et guider dans l'adéquation aux nouvelles demandes, aux priorités de l'**Agenda** de R&D et aux nouveaux programmes d'incorporation et de formation de ressources humaines ;
4. Le Séminaire "Stratégies d'Intelligence Compétitive - Études de Cas et Expériences", le 15 mai 2002, à Brasília, Brésil. Ont été conférenciers le professeur Luc Quoniam, de l'Université du Sud, Toulon-Var, sur "Des projets d'intelligence compétitive en brevets réalisés au Brésil et en France" et le professeur Leandro Innocentini Lopes de Faria, chercheur du Noyau d'Informations Technologiques en de Matériels de l'Université Fédérale de Sao Carlos (NIT Matériels) qui a présenté une étude de cas de prospection technologique et a avancé quelques résultats préliminaires de l'analyse bibliométrique de la base de données **Collection documentaire** de l'Embrapa.
5. Le Séminaire "Technologies pour le Traitement d'Information à Embrapa", du 10 au 14 décembre 2001, à Brasília, Brésil. Ont été conférenciers le professeur Henri Dou, de l'Université Aix Marseille III, le professeur Luc Quoniam et le professeur Eric Boutin, de l'Université du Sud, Toulon-Var. Ont été entraînés 30 techniciens du Siège de l'Embrapa et des centres de recherche de l'Amazonie (Embrapa Amazonie Orientale) et du District Fédéral. Ce séminaire a permis à l'Embrapa de commencer à travailler efficacement avec le traitement et l'analyse d'informations. Elle a reçu, pour utilisation interne, les logiciels Dataview et Matrisme, de l'Université Aix Marseille III.

En 13 mars, 2006 il a été réalisé le "II Séminaire International Outils pour l'Intelligence Compétitive", à Brasília, Brésil. Parmi les conférenciers, le professeur Alan Porter, de l'Université de Georgia Tech, sur "L'exploration de textes en Science et Technologie" ; Webb Myers, directeur de relations avec le client de la **Search Technologies**, sur "Le logiciel de text-mining VantagePoint" ; Roberto Pentead, doctorant de l'Université du Sud, Toulon-Var, sur "Médiométrie : L'analyse de la performance de l'Embrapa dans les médias avec l'aide de méthodologies et logiciels bibliométriques".

Le projet "Développement de bases de données"

Le sous-projet 1. Création d'un système pilote d'intelligence compétitive et de gestion de la connaissance pour la communication à l'Embrapa a eu comme prédécesseur un projet initié en 1997 appelé "Développement de base de données de la communication", créé par le Bureau de Communication Sociale (ACS) de l'Embrapa pour automatiser et rationaliser les services

aux publics interne et externe. Ce projet aussi a été conduit par le doctorant et a reçu le Prix National de l'Embrapa en 1997, dans la catégorie "Analyse et Amélioration de Processus".

Le projet Développement de base de données de la communication, d'un côté fait attention au public interne avec la création d'un Clipping électronique, ce qui a expédié la circulation et la recherche d'informations stratégiques par les centres de recherche. De l'autre côté, le projet a fait attention au public externe avec la création d'une Banque de nouvelles, pour expédier la fourniture d'informations aux journalistes et aux autres intéressées en recevoir des nouvelles de l'Embrapa.

Avec la Banque de nouvelles de l'ACS disponible sur l'Internet en <http://www21.sede.embrapa.br/noticias/banco_de_noticias/index_htm>, se sont réunies toutes les nouvelles écrites par les journalistes de l'ACS et des 37 centres de recherche dans un seul dépôt, ce qui a permis une récupération rapide des données et un rapide service aux clients externes. Le service à la presse ou à des consultations du public est devenu plus qualifié et plus rapide.

Tant la Banque de nouvelles comme le Clipping électronique ont été créés sur l'Internet, en utilisant le langage HTML. Les deux ont évolué pour le logiciel Lotus Notes¹⁸, la Banque de Nouvelles en 1997 et le Clipping Électronique, en 2002. À partir de 2003, l'Embrapa a adopté une politique d'appui au logiciel libre et en conséquence la Banque de Nouvelles a passé à fonctionner avec le logiciel Zope/Plone¹⁹.

¹⁸ Disponible chez IBM en "<http://www-306.ibm.com/software/lotus/>"

¹⁹ Disponible en "<http://www.contentmanagementsoftware.info/>"

Audit de la collection documentaire et analyse de la production scientifique

L'Ainfo et son histoire

Lenk (2000) indique que l'Embrapa a créé en 1991, le logiciel Ainfo pour organiser, récupérer et disséminer les informations techniques et scientifiques produites et/ou acquises par l'entreprise.

Initialement, les efforts ont été concentrés sur la formation de bases de données bibliographiques : Collection de Documents - livres, brochures, separata, articles de périodiques, travaux présentés dans des congrès, logiciels, cassettes vidéos, thèses, autres matériels spéciaux - et la collection de périodiques. Plus tard, avec la mise en oeuvre de relations entre bases de données pour automatiser routines de prêt des travaux ont été initiés avec des bases de données d'institutions et des ressources humaines. Le contrôle des processus bibliographiques acquisition, commutation et service de SDI, s'est automatisé aussi, avec le transfert de données entre les bases (p.2).

Toutes les bases de données bibliographiques des centres de recherche de l'Embrapa ont été réunies dans l'Ainfo. Le logiciel a été constamment perfectionné et actuellement permet l'exportation des données pour le format MARC 21 et il suit aussi la norme Z 39.50, tous les deux utilisés par les grandes bibliothèques du monde.

Cette réunion des données bibliographiques de tous les centres de recherche a formé une collection appelée "Base de données de la recherche agronomique - BDPA" (Embrapa, 2006a) et aujourd'hui est mise sur l'Internet, pour consultation publique. L'adresse est <http://www21.sede.embrapa.br/bibliotecas_virtuais/index_htm>.

Avec la BDPA, l'Embrapa viabilise l'accès à la littérature technique et scientifique produite et réunie dans ses centres de recherche, dans tout le territoire national. Elle se compose des bases suivantes : a) Collection documentaire de l'Embrapa ; b) Production scientifique de l'Embrapa ; c) Catalogue collectif de périodiques de l'Embrapa ; d) Cadastre d'institutions. Cet effort actuellement est coordonné par l'Embrapa Information Technologique (SCT) et opère en deux instances : La Commission permanente pour l'Ainfo (CP-Ainfo) chargée de définir des critères et de la méthodologie pour l'évolution du logiciel et le Groupe de travail de référence bibliographique (GT-Refer) qui s'occupe des normes de référence bibliographique à l'Embrapa, réunies dans le Manuel d'information et de documentation (Embrapa, 2001).

Les bases Embrapa fonctionnent en parallèle avec les bibliothèques. Trois centres de recherche, l'Embrapa Café (CAFE), l'Embrapa Trigo (CNPT) et l'Embrapa Élevage du Bétail Sud (CPPSUL) ne sont pas rapportés dans cette liste. Le premier n'a pas de bibliothèque et les deux autres, pour n'utiliser pas l'Ainfo. A remarquer encore qu'un centre de recherche, l'Embrapa Moyen-Nord (CPAMN) apparaît deux fois (CPAMN et CPAMN-UEPP) parce qu'il utilise deux bibliothèques.

La procédure de jonction des bases bibliographiques des 39 bibliothèques de l'Embrapa a créé un problème pour les administrateurs de la base : Comment traiter les doublons des diverses collections ? La question a été résolue par la création d'un algorithme pour identification des doublons qui compare le titre, l'auteur, l'année de publication et avec l'application de normes de normalisation et de remplissage des champs.

Entre 1996 et 2005, le nombre de références de BDPA s'est accru en 111,48%, passant de 245.622 à 519.453 références. De ce total, l'algorithme d'identification de doublons (archives avec le même titre, auteur et année de publication) a identifié 123.778 références, environ 23,8% du total. Ceci équivaut à 395.675 références sans duplication. La relation est illustrée au tableau 5.

Tableau 5 : Base Collection documentaire ; Nombre de références indexées 1996/2005

Centre	1996 Avr	1997 Jan	1997 Jul	1998 Jan	1998 Nov	1999 Avr	2000 Avr	2000 Nov	2001 Mai	2001 Nov	2002 Nov	2003 Avr	2003 Nov	2004 Avr	2004 Nov	2005 Mai
AI/SEDE	8.373	9.158	11.311	13.013	14.195	16.128	17.302	18.770	19.283	20.164	22.836	23.235	23.757	24.602	26.993	28.849
CENARGEN	7.246	7.817	7.817	9.997	11.354	12.501	14.661	16.730	16.972	17.123	21.018	21.467	21.748	21.936	22.149	22.503
CNPA	11.853	11.847	12.063	12.238	12.503	12.787	13.489	13.489	14.104	14.297	14.932	15.010	15.369	15.641	15.820	16.120
CNPA B	22.431	22.549	22.575	22.881	22.967	23.124	23.259	23.337	23.539	23.746	26.404	26.404	26.947	28.018	29.174	30.588
CNPA F	11.520	11.584	11.771	12.403	12.887	13.295	15.517	16.153	16.237	16.774	18.237	17.860	19.373	19.991	20.554	21.395
CNPAT	2.894	3.403	4.084	4.286	4.286	4.566	4.620	5.187	5.496	5.496	6.505	6.806	7.246	7.340	7.576	8.052
CNPC	5.620	8.994	8.994	9.978	9.978	11.775	12.386	12.765	13.141	13.742	14.571	14.861	15.580	15.956	16.325	16.664
CNP DIA	1.907	2.067	2.121	2.261	2.496	2.503	5.838	5.574	5.669	5.669	6.678	5.289	5.893	3.459	3.679	3.997
CNPF	16.431	17.074	17.075	18.128	18.462	19.434	22.142	23.746	25.001	26.533	29.407	29.782	27.450	28.498	29.553	30.588
CNPGC	3.505	3.619	3.820	4.533	4.691	5.017	6.135	6.892	7.241	7.487	8.262	8.461	8.708	8.805	8.940	9.112
CNPG L	11.358	11.430	11.617	11.777	11.821	12.156	12.709	12.985	13.124	13.124	13.494	13.494	13.508	13.508	13.526	13.543
CNP H	6.400	7.179	8.556	9.441	11.721	12.786	17.429	19.337	20.162	20.719	23.051	23.865	25.358	26.267	27.458	28.722
CNPM	446	463	463	581	583	656	1.032	1.142	1.260	1.260	1.340	74	613	720	1.175	1.173
CNPM A	1.971	2.822	2.890	2.937	3.094	3.202	3.303	3.303	2.747	2.767	19.090	3.763	3.897	4.006	4.115	4.261
CNPM F	7.504	7.740	8.393	9.422	10.516	11.849	14.368	17.287	18.377	18.377	19.019	19.101	19.101	19.307	19.730	19.835
CNPM S	1.094	2.327	3.904	6.434	8.389	10.092	11.521	12.298	12.764	12.874	14.308	14.449	15.002	15.214	16.012	16.413
CNPS	3.234	3.359	3.859	6.645	7.388	7.962	8.972	8.972	9.020	9.947	10.493	10.530	10.583	10.666	10.789	10.998
CNPS A	2.954	3.173	3.989	4.235	4.495	7.572	8.447	9.264	9.981	11.156	11.979	12.210	12.814	13.052	13.909	14.098
CNPS O	12.958	13.209	13.209	13.434	13.239	13.776	14.614	15.645	16.360	17.213	18.106	18.477	21.096	22.369	22.956	23.375
CNPT	10.375	10.954	14.843	15.368	15.368	16.117	17.268	17.268	17.268	17.268	-	-	-	-	-	-
CNPT IA	5.080	5.550	5.982	7.237	7.902	8.073	7.939	8.002	7.597	7.686	7.828	7.927	8.294	8.590	8.518	8.552
CNPU V	5.895	6.782	6.846	6.902	6.920	6.942	4.939	4.939	4.840	4.840	4.363	4.498	4.589	4.747	5.175	5.683
CPAA	206	529	888	1.394	1.905	2.314	5.741	4.508	5.222	5.222	6.497	6.954	7.384	8.198	9.202	10.377
CPAC	5.934	6.596	8.587	10.754	12.004	13.468	16.968	18.172	18.504	18.957	20.965	21.796	22.302	22.739	23.150	24.192
CPACT-Sede	5.085	5.099	5.198	4.548	4.556	4.649	4.737	8.026	8.026	8.298	6.642	6.647	7.350	7.512	7.584	7.704
CPACT-TB	3.155	3.232	3.155	3.197	3.208	3.233	3.233	-	-	-	-	-	-	-	-	-
CPAF-AC	89	89	362	561	576	714	1.504	2.379	2.424	2.592	2.639	2.652	3.846	5.107	7.163	8.498
CPAF-AP	924	1.023	1.280	1.560	1.879	2.220	3.710	4.671	5.347	5.595	6.150	6.150	6.487	6.487	6.869	7.325
CPAF-RO	588	588	1.403	2.292	2.650	3.753	4.859	4.859	4.859	4.859	5.726	5.839	6.891	7.046	7.270	8.664
CPAF-RR	1.439	1.534	2.311	2.385	2.821	2.990	3.057	3.586	4.015	4.326	-	4.797	4.847	4.848	4.939	5.303
CPAMN-Sede	1.345	1.942	2.966	3.355	4.088	5.161	7.777	7.777	11.001	13.485	15.188	15.835	16.621	16.815	17.181	17.692
CPAMN-UEPP	6.856	7.205	7.587	7.709	7.955	7.649	8.787	8.787	9.331	9.578	9.949	10.050	10.260	10.416	10.535	10.693
CPAO	11.210	11.605	12.183	12.430	12.513	12.986	13.839	14.466	15.059	15.631	16.400	16.546	18.290	19.221	21.049	21.964
CPAP	5.343	5.885	6.986	7.598	8.237	9.052	15.051	19.065	21.282	21.282	25.487	25.696	26.060	26.229	26.573	26.908
CPATC	9.877	10.158	10.603	10.743	14.105	16.024	16.860	16.860	16.860	16.860	-	17.332	17.487	17.722	17.964	18.066
CPATSA	12.505	12.661	13.493	14.319	14.863	15.676	17.569	18.525	19.103	20.115	23.391	24.056	24.928	25.893	27.131	28.326
CPATU	9.039	9.310	10.053	11.182	12.030	13.604	19.036	20.863	22.577	24.590	-	28.970	30.388	31.599	32.118	32.749
CPPSE	18.670	19.197	19.497	19.617	19.676	19.672	19.987	20.396	20.548	13.502	13.633	13.766	14.041	14.290	14.418	14.627
CPPSUL	8.439	8.466	8.266	8.266	8.359	8.603	8.961	8.961	9.286	9.286	-	-	-	-	-	-
CTAA	2.757	2.794	2.800	2.889	4.148	4.322	4.744	4.744	4.982	4.982	5.789	5.812	5.886	5.976	6.040	6.147
Total reg.	264.510	281.013	303.800	328.930	350.828	378.403	434.310	459.730	478.609	487.422	470.377	504.311	529.994	542.790	563.312	583.756
Reg. Valides	245.622	251.152	270.426	284.387	304.362	321.564	368.159	390.984	409.831	425.224	391.635	445.186	466.088	480.582	499.780	519.453
Sans duplic.	-	-	231.810	241.685	254.805	266.271	296.939	315.316	328.534	338.604	321.478	358.275	374.613	386.695	383.698	395.675
% duplicata	-	-	14,27	15,01	16,28	17,19	19,34	19,35	19,83	20,37	17,91	19,52	19,62	19,53	23,22	23,82

Parmi les bases de BDPA, celle de la Production scientifique de l'Embrapa réunit les documents édités par l'entreprise et par autres organismes et les travaux, articles, communications, thèses et dissertations publiées par ou en collaboration par ses employés. En 2005, cette base avait 72.015 références valides. Ses détails sont fournis au tableau 6.

Tableau 6 : Base Production scientifique ; Références valides, avril 2005

Centres	total reg.	total valides	prod. scient.	prod. scient. valides	% valides
CPATU	32749	31543	4855	4750	15
CNPAB	30588	28691	1918	770	2,7
CNPF	30588	28428	3087	3055	10,7
CPATSA	28326	28183	3905	3881	13,8
AI-SEDE	29017	27377	8189	8070	29,5
CPAP	26908	26082	1530	1503	5,8
CNPH	28722	23943	4939	4731	19,7
CPAC	24192	22883	5223	5094	22,2
CNPSO	23375	22824	7939	7884	34,5
CPAO	21964	21004	1047	1041	4,9
CENARGEN	22503	19318	1938	1902	9,8
CNPMF	19835	18956	269	195	1
CPATC	18066	17678	527	525	2,9
CPAMN	17692	17497	1946	1928	11
CNPMS	16413	15952	5658	5379	33,7
CPPSE	14627	14622	2565	2565	17,5
CNPA	16120	13526	1283	1243	9,2
CNPAF	21395	13131	2373	2349	17,9
CPAMN-UEPP	10693	10626	152	151	1,4
CNPGL	13543	10105	181	177	1,7
CPAA	10377	9614	1935	1732	18
CPAF-RO	8664	8343	790	784	9,4
CPAF-AC	8498	8242	840	805	9,7
CNPC	16664	7989	1367	1361	17
CNPS	10998	7872	1110	1099	13,9
CNPGC	9112	7713	1951	1716	22,2
CPACT	7704	7644	684	684	8,9
CNPAT	8052	7631	321	311	4,1
CPAF-AP	7325	7234	357	355	4,9
CNPTIA	8552	7031	1055	1000	14,2
CNPSA	14098	6147	2250	715	11,6
CNPUV	5683	5667	213	213	3,7
CTAA	6147	4439	1375	1340	30,2
CNPDIA	3997	3799	1357	1308	34,4
CNPMMA	4261	3637	793	759	20,8
CPAF-RR	5303	3086	267	267	8,6
CNPM	1173	1164	382	373	32
CNPT	0	0	0	0	0
CPPSUL	0	0	0	0	0
Total	583924	519621	76571	72015	13,8

Note. total reg. = nombre total de références dans la Collection Documentaire.
total valides = nombre de références marquées pour apparaître dans la Collection Documentaire.
prod. scient. = nombre de références concernant à des documents de la Production scientifique de l'unité de recherche de l'Embrapa.
prod. scient. valides = nombre de références marquées pour apparaître dans la Production scientifique.
% valides = pourcentage prod. scient. valides x total valides.

Dans un effort pour rendre disponible d'une manière massive la connaissance scientifique disponible sur l'agriculture, l'élevage et la forêt tropicale et tempérée, les centres de recherche de l'Embrapa ont commencé à digitaliser et à mettre sur l'Internet par moyen de l'Ainfo, leurs travaux scientifiques les plus importants. En 2005, 7% de la Collection documentaire de l'Embrapa avait déjà été digitalisée et était disponible en ligne. Quelques centres ont créé des vraies bibliothèques digitales avec toutes leurs publications. Ceci est le cas de la Bibliothèque Électronique de l'Embrapa Cerrados (Embrapa, 2006b), disponible en ligne <http://bbeletronica.cpac.embrapa.br/versaoframe/principal_02.htm>, qui réunit le texte intégral dans le format PDF (Adobe Acrobat) des publications techniques et scientifiques qui y ont été produites.

L'audit de la collection documentaire

Entre novembre 2001 et février 2002, un audit a été réalisé à la Collection documentaire de l'Embrapa par la bibliothécaire Vera Lucia de Campos Octaviano, ayant comme objectif "d'améliorer la qualité" de la BDPA par "l'identification des doublons" ; "La localisation d'erreurs ou d'inconsistances qui ont causé cette duplication" ; "La notification aux bibliothèques de l'Embrapa en sollicitant d'effectuer les corrections" et "l'évaluation des corrections effectuées par les bibliothèques de l'Embrapa" (Octaviano, 2002, p.1). L'audit s'est justifié par le fait que l'Association Brésilienne des Normes Techniques (ABNT) avait modifié, en août 2000, la norme brésilienne de références bibliographiques et le Manuel d'information et de documentation (Embrapa, 2001) avait aussi été modernisée, en prenant comme base cette nouvelle norme, ABNT NBR 6023 - 2000. L'occasion était donc, propice pour une vérification de l'état général de la base. Les champs audités ont été "auteur", "titre", "source" et "année de publication".

Dans le champ "source", les problèmes suivants ont été remarqués : Tous les éléments écrits en majuscules et avec un point, après l'éditeur, absence des indicatifs de l'état pour des locaux homonymes et du District Fédéral (solution = identifier le registre pour correction manuelle) ; Des données non appartenant au champ, de l'inconsistance dans la forme d'indiquer le lieu,

éditeur et date non-identifiés, écriture incorrecte de l'éditeur, deux-points après le lieu dans des entrées par auteur corporatif (solution = substituer les formes incorrectes par les formes correctes), espace avant et/ou après le lieu (solution = retirer l'espace). Octaviano affirme encore que des références ont été détectées avec "absence de données dans tous les champs évalués" (2002, p.8), des références vides qui n'avaient pas été effacés.

Selon Octaviano (2002), outre les erreurs de dactylographie qui "ont été remarquées dans tous les champs évalués" (p.4) un grand volume de références avait des problèmes d'adéquation à la nouvelle norme ABNT NBR 6023. Il manquait, par exemple, "un espace après le point de l'abréviation du (des) prénom (s) du (des) auteur (s), du volume, du nombre (fascicule) et de la (des) page (s)" (p.4) :

RAMEY JUNIOR, H.H.; TURNER JUNIOR, J.H.; WORLEY JUNIOR, S.

Ce champ auteur devrait être corrigé pour :

RAMEY JUNIOR, H. H.; TURNER JUNIOR, J. H.; WORLEY JUNIOR, S.

(et)

v.21, n.3/4, p.219-224

Ces informations devraient être corrigées pour :

v. 21, n. 3/4, p. 219-224

Octaviano a détecté et documenté plusieurs d'autres inconsistances dans les champs audités qui seront mentionnées en détail plus tard au chapitre "Le nettoyage de la Collection documentaire et de la Production scientifique". En général, elle affirme :

en se prenant, par exemple, les premières archives évaluées, totalisant 200 pages, 67 ont été transférées pour l'archive des Doublons (33.5%), 11 pour ce des Corrects (5,5%) et 122 pages pour l'archive des Inconsistantes (formé par les références non-dupliqués, mais avec des erreurs de dactylographie, des inconsistances dans l'indexation et dans le remplissage des champs, ce qui correspond à 61% des références (Octaviano, 2002, p.78).

Octaviano mentionne que comme "la quantité d'erreurs trouvée à été beaucoup plus grande que celle prévue" a été déterminée par la CP-Ainfo que les phases finales de l'audit, la notification des erreurs aux bibliothèques et l'évaluation des corrections, seraient réalisées après l'adoption de mesures de rectification par la CP-Ainfo. Ainsi, affirme Octaviano, "cet

audit servira comme base pour l'adoption de providences" (2002, pp.78-79) ayant comme but la "réduction du nombre de doublons et l'amélioration de la qualité de la base de données" (2002, p.1).

En comptant que cet échantillon de 200 pages est équivalent à 1,3% du total de pages de la base (15.603 pages), une solution manuelle comme celle imaginée, en comptant avec le dévouement total et exclusif d'une bibliothécaire et en supposant qu'elle emploie, en moyenne, 10 minutes pour réviser et corriger chacun des 4.160 références (94.5%) des duplicata ou inconsistants, mobilisera 694 heures ou pratiquement 90 jours de service (de 8 heures). En considérant encore un salaire moyen mensuel de R\$ 3.000,00 (E\$ 1.200,00) d'un professionnel de la Science de l'Information, le coût de ce travail atteindrait à R\$ 12.000,00 (E\$ 4.800,00 environs).

Cette projection pour 1.3% de la base donne une idée du travail et du coût pour corriger tous les 319.980 références de la base à la main. Dans des termes bruts seraient nécessaires 25 ans et quatre mois de service exclusif de cette bibliothécaire, à un coût pour l'entreprise de R\$ 909.000,00 (E\$ 363.600,00). Outre les ressources financières il manquerait aussi du temps et du personnel dans les centres de recherche pour se consacrer à cette tâche. Ainsi, avec le rejet dans la pratique de la solution manuelle, la CP-Ainfo, avec l'orientation et l'assistance des Universités d'Aix Marseille III, du Sud, Toulon-Var et Fédérale de Sao Carlos et par l'initiative du Projet intelligence compétitive et gestion de la connaissance à l'Embrapa, a commencé à étudier une solution automatique en utilisant des méthodes bibliométriques pour reformatage et traitement de textes et références, en particulier, avec le logiciel Infotrans.

Une solution bibliométrique pour la création d'indicateurs

Dans le travail intitulé "Application de la bibliométrie dans la construction d'indicateurs sur la production scientifique de l'Embrapa", Penteado et al. (2002) ont proposé une première solution intégrée concernant les problèmes de la Collection documentaire et la question de la nécessaire automatisation des opérations de construction d'indicateurs de production scientifique à l'Embrapa par l'application de la scientométrie. Ils indiquent :

DES INFORMATIONS DE GESTION

L'Embrapa dispose de bases de données de gestion, entre autres, le système plan annuel de travail (SISPAT), le système d'évaluation des unités (SAU) et le système

d'évaluation et de récompense par des résultats (SAPRE) qui administrent une bonne partie des processus, plans, activités et résultats.

Les informations du SISPAT sont utilisées pour la procédure de négociation, d'accompagnement et d'évaluation des objectifs du SAPRE, ainsi que pour l'élaboration de rapports de gestion pour le Ministère de l'Agriculture et le Tribunal de Comptes de l'Union, entre autres, outre de servir de base pour la procédure de veille institutionnelle et de planification stratégique de l'entreprise. Celle-ci utilise aussi les indicateurs de performance des unités de travail pour évaluer et récompenser les employés et les équipes de travail selon les objectifs de productivité de l'entreprise comme un tout.

Les principaux indicateurs calculés par le SISPAT/SAU/SAPRE sont :

Production Technique et Scientifique : 1.1-Article en Périodique Indexé ; 1.2-Chapitre en Livre Technique et Scientifique ; 1.3-Article dans des Actes de Congrès / Note Technique ; 1.4-Résumé dans des Actes de Congrès ; 1.5-Orientation de Thèse de Doctorat ;

Production de Publications Techniques : 2.1-Circulaire Technique ; 2.2-Comunicat Technique ; 2.3-Bulletin de Recherche et Développement ; 2.4-Documents ; 2.5-Organisation/Édition de Livres ou de Système de Production ; 2.6-Article de Divulgateur dans les Mass Médias ;

Développement de Technologies, Produits et Services : 3.1-Cultivar Produite/Lancée et Événement Élite ; 3.2-Cultivar Expérimentée/Recommandée ; 3.3-Pratique/Processus Agricole ; 3.4-Race-Type ; 3.5-Sous-produits Agricoles ; 3.6-Processus Agro-industriel ; 3.7-Méthodologie Scientifique ; 3.8-Machine, Équipement ou Installation ; 3.9-Lignées ; 3.10-Surveillance/Zoning ; 3.11-Logiciel ; 3.12-Organisation d'Informations pour Base de Données.

INFORMATIONS BIBLIOGRAPHIQUES

Connaître en détail sa production scientifique et technologique est fondamental pour qu'une organisation de recherche, puisse orienter sa planification stratégique et évaluer si les actions de recherche sont conformes aux programmes de recherche et développement.

L'Embrapa compte avec une collection décentralisée, disponible dans les bibliothèques de ses plusieurs unités. Incorporé à la Base de données de la recherche agronomique, l'entreprise produit la base Collection documentaire (disponible sur CÉDÉROM ou à travers l'Internet dans le site <http://www.cnptia.embrapa.br/servicos/bdpa/frame_bdpa.html>) qui réunit environ 340.000 références bibliographiques et pratiquement la totalité des travaux scientifiques et technologiques des chercheurs et en plus, les publications présentes dans les bibliothèques de l'entreprise. La Collection documentaire peut être utilisée pour la récupération de données sur les publications de l'Embrapa,

faisant partie du processus d'analyse et de construction d'indicateurs. Néanmoins, le volume d'informations disponibles et la finalité de la base créée pour récupérer des publications, deviennent des défis pour ce processus d'analyse de la production technique et scientifique.

Selon Faria (2001), la disposition de bases de données électroniques a augmenté l'offre d'informations sur pratiquement tous les sujets (Dou, 1994, Porter, 2003, Wormell, 1998). Mais, quand il y a beaucoup d'informations, comme c'est le cas dans la Collection documentaire de l'Embrapa, le processus d'analyse et de synthèse des informations est plus complexe, exigeant l'utilisation de méthodologies appropriées. Les bases de données offrent des informations abondantes, structurées et électroniques. Ceci a motivé le bourgeonnement de méthodologies d'analyse d'informations assistées par ordinateurs, visant à l'élaboration d'informations à haute valeur ajoutée pour les Décideurs (Quoniam, 1992).

Ce travail vise à la réalisation d'une analyse de la production scientifique de l'Embrapa pour vérifier le volume de la production scientifique et technologique annuelle, les sujets de recherche abordés par l'entreprise et l'évolution de ces sujets au long de son histoire. On cherchera aussi à créer des indicateurs sur la production scientifique des chercheurs de l'Embrapa et de ses centres de recherche.

MÉTHODOLOGIE

Les logiciels Infotrans (de l'IUK GmbH, d'Allemagne) pour le reformatage des données, Dataview et Matrisme (du CRRM, Université d'Aix Marseille III) seront utilisées pour le traitement bibliométrique des données, Microsoft Excel pour l'élaboration de graphiques et des tableaux (...). La Collection documentaire de l'Embrapa est la base utilisée.

En préliminaire trois indicateurs généraux ont été recherchés : Production scientifique de l'Embrapa par année, production scientifique de l'Embrapa par chercheur et production scientifique de l'Embrapa par centre de recherche. Les champs de la Collection travaillés ont été : Auteur, origine du matériel, affiliation, année de publication et titre.

Une consultation réalisée avec la bibliothécaire Vera Lúcia de Campos Octaviano a identifié les principales inconsistances des références, dans les champs "auteur" et "titre". Par exemple : Entrées comme "João José Silva", "Silva, João José", "Silva, J. J.", "Silva, JJ" et autres variantes sont normalisés pour "Silva, J. J." Les champs "origine du matériel" et "affiliation" avec des unités de recherche écrites dans des différentes façons sont normalisés. Les champs sélectionnés sont traités avec le logiciel Infotrans lequel permet la création de règles génériques de normalisation pour tous les noms, titres et mots reformatés. D'autre part, le logiciel élimine automatiquement les duplications de références.

À l'issue de cette étape, la base est prête pour l'analyse bibliométrique, les champs extraits (auteur, origine du matériel, affiliation, année de publication et titre) sont traités dans Dataview, avec la génération des matrices; Année de publication-auteur, année de publication-affiliation, auteur-titre, affiliation-titre, auteur-affiliation. Ces matrices sont exportées pour Matrisme, visant la création de réseaux relationnels d'auteurs et des thèmes et pour MS Excel, pour l'élaboration de graphiques et de tableaux. (...)

DISCUSSION

L'importance et la valeur des indicateurs reposent sur la possibilité de quantifier des "choses" intangibles comme la génération de connaissances. Voici quelques exemples d'indicateurs scientifiques : Nombre de publications par an de la société, nombre de chercheurs actifs, évolution des principaux thèmes de recherche an par an, ressources financières appliquées et bourses accordées (...)

Faria (2001, p. 9) ajoute que "les indicateurs sont une façon de synthétiser et d'ajouter de la valeur aux informations". Ils synthétisent le processus de traitement automatisé des informations comme dans la figure suivante.



Note. [De gauche à droite, de haut en bas : Récupérer des données, traitement bibliométrique, traitement statistique, représentation graphique. Grande quantité de données, analyse difficile ---> Des indicateurs de haute valeur agrégée.]

Figure 3 : Le cycle du traitement automatique de l'information (Penteado et al., 2002, p.8)

Penteado et al. (2002, p.8) soulignent "qu'il est toujours nécessaire d'améliorer les données et aussi d'insérer des données externes ou créer des nouvelles données pour les étudier en profondeur". Et continuent :

La Collection documentaire de l'Embrapa même en réunissant pratiquement la totalité des travaux des chercheurs de l'entreprise, n'a pas été développée pour permettre l'analyse des données, mais au contraire, pour identifier et localiser rapidement les documents. Son utilisation par l'analyse automatique exige, pourtant, une longue étape de préparation et reformatage des données, sans laquelle il n'y aurait pas de qualité dans les informations. (...)

Le processus continu de transformation des données brutes en connaissance stratégique, s'est accru exponentiellement de nos jours. La nécessité même des entreprises d'obtenir connaissance de leur chaîne de production a conduit à une utilisation intensive des technologies de l'information mises à la disposition par l'industrie de logiciels dans tout le monde.

Comme il a été déjà dénommé à ce travail, dans la majorité absolue des cas, les données disponibles dans les bases de données des organisations ne sont pas prêtes pour le processus d'analyse, une difficulté supplémentaire du processus de transformation d'information en connaissance.

L'objet de la technologie des supermarchés de données (DW) est de permettre l'organisation de ces données, de les traiter et de les utiliser d'une façon efficace et encore, faciliter leur récupération pour la prise de décision parmi une interface amiable et de facile accès. Les logiciels Infotrans et Dataview qui exécutent le travail de nettoyage et analyse des données, intègrent le travail de DW. Cette technologie renvoie à la nécessité d'avoir une banque d'informations centrale et intégrée qui pourra devenir et composer de nouvelles structures de données plus tard.

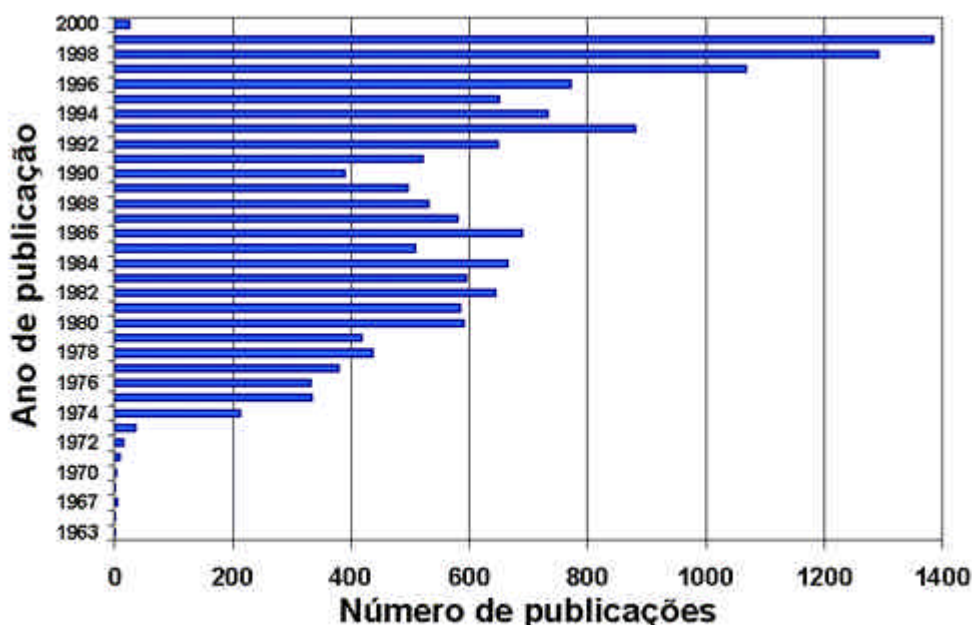
Selon Inmon et Hackathorn (1994) la majorité des outils dans le marché des logiciels offrent une manière de filtrer des données pour garantir leur qualité pendant le processus d'extraction et transformation. Des outils spécifiques comme Infotrans, Dataview et Matrisme offrent des mécanismes bien plus sophistiqués en conséquence une plus grande crédibilité à l'information. Par exemple : Dans le cas de la base de données SISPAT, les indicateurs sont générés à partir d'informations réunies par les chercheurs et techniciens des centres de recherche de l'Embrapa. Quand ces données sont réunies d'une façon automatique, dans une base publique, nous évitons la subjectivité du travail humain.

RÉSULTATS

Comme résultat de la reformatage et du traitement bibliométrique des données extraites de la Collection documentaire ont été automatiquement produits, en phase pilote, trois indicateurs scientifiques : Classifications, tableaux et figures de la production scientifique de l'Embrapa.

Le premier, un indicateur de fréquence basique comme l'évolution du nombre de publications, représenté dans la figure 4. Nous pouvons observer que, depuis 1997, la production de publications par les chercheurs de l'Embrapa a changé de palier, en dépassant les 1.100 par an et s'en accroissant significativement après.

Note. [De gauche à droite : Année de publication x nombre de publications. Source Collection documentaire de l'Embrapa, jusqu'à 2000.]



Fonte: Acervo Documental Embrapa, até 2000

Figure 4 : Production scientifique de l'Embrapa par année de publication (Penteado et al., 2002, p.10)

Le deuxième indicateur produit est celui de la production scientifique des chercheurs de l'Embrapa (figure 5), avant et après le reformatage. Après le traitement préliminaire des données, considérant la production des 29 ans de l'entreprise, un résultat initial est obtenu : quatre chercheurs, Santos, H.P.S.; Campos, O.F.; Souza, C.N.A.; Ignaczak, J.C., ont produit de 257 à 150 travaux.

Note. [De gauche à droite : Avant le reformatage x après le reformatage. Source Collection documentaire de l'Embrapa, jusqu'à 2000.]

Antes da reformatação		Após a reformatação	
1 Pesquisadores	Publicações	1 Pesquisadores	Publicações
2 EMERAPA (BRASILIA, DF)	979	2 SANTOS, H.P.S.	257
3 EMERAPA CENTRO DE PESQUISA AGROPECUARIA DOS CERRADO	432	3 CAMPOS, O.F.	157
4 EMERAPA	252	4 SOUSA, C.N.A.	155
5 EMERAPA SERVIÇO NACIONAL DE FAVANTAMENTO E CONSERVA	206	5 IGNAZAK, J.C.	150
6 EMERAPA CENTRO NACIONAL DE PESQUISA DE TRIGO (PASSO FL	173	6 ARIAS, G.	139
7 EMERAPA CENTRO NACIONAL DE PESQUISA DE SOJA (LONDRINA	169	7 MOREIRA, J.C.S.	138
8 EMERAPA DEPARTAMENTO DE INFORMACAO E DOCUMENTACAO	152	8 CUNHA, G.R.	137
9 SANTOS, H.P. DOS	144	9 ALVIM, M.J.	137
10 ALVIM, M.	134	10 BAIER, A.C.	136
11 ROSA, A. DO N.	121	11 ROSA, A.N.	127
12 EUCLIDES FILHO, K.	117	12 AMEROSI, I.	125
13 EMERAPA DEPARTAMENTO DE DIFUSAO DE TECNOLOGIA BRASIL	116	13 MINELLA, E.	123
14 EMERAPA (BRASILIA, DF)	115	14 LHAMBY, J.C.B.	118
15 TEODORO R.L.	104	15 EUCLIDESFILHO, K.	118
16 POTT, A.	103	16 BOTREL, M.A.	111
17 VALÉRIO, J.R.	100	17 BORATO, E.R.	109
18 EMERAPA DEPARTAMENTO DE DIRETRIZES E METODOS DE PLAN	100	18 TEODORO, R.L.	108
19 BIANCHIN, I.	97	19 VERNEQUE, R.S.	106
20 VILELA, D.	94	20 POTT, A.	106
21 MARTINEZ M.L.	94	21 PICININI, E.C.	104
22 ARIAS, G.	89	22 LEMUS, A.M.	104

Fonte: Acervo Documental Embrapa, até 2000

Figure 5 : Production scientifique de l'Embrapa par chercheur (Penteado et al., 2002, p.11)

Enfin, le troisième indicateur concerne la production scientifique dans les centres de recherche de l'Embrapa (figure 6), avant et après le reformatage. Même si ces résultats préliminaires peuvent être affinés, les centres les plus productifs dans ces 29 ans sont le Centre National de Recherche de Blé, le Centre National de Recherche de Bovins à Lait et le Centre National de Recherche de Bovins à Viande, qui ont produit de 2459 à 1735 travaux.

Note. [De gauche à droite : Avant le reformatage x après le reformatage. Source Collection documentaire de l'Embrapa, jusqu'à 2000.]

Antes da reformatação		Após a reformatação			
1	Instituição	Publicações	1	Instituição	Publicações
2	CAMPO VAZIO	6269	2	EMBRAPA	2589
3	CNPGL	1827	3	EMBRAPA - CNPT	2459
4	EMBRAPA. CENTRO NACIONAL DE PESQUISA DE GADO DE CORTE (CAN)	1183	4	EMBRAPA - CNPGL	2172
5	FMRRAPA	776	5	FMRRAPA - CNPSC	1735
6	EMBRAPA. CENTRO DE PESQUISA AGROPECUARIA DO PANTANAL (CO)	288	6	EMBRAPA - CPAP	664
7	EMBRAPA. CENTRO DE PESQUISA AGROPECUARIA DO PANTANAL (CO)	259	7	EMBRAPA - CPAC	634
8	EMBRAPA - CNPT, PASSO FUNDO, RS.	247	8	EMBRAPA - CPATSA	568
9	CNPQ	196	9	EMBRAPA - CNPSO	432
10	EMBRAPA. CENTRO NACIONAL DE PESQUISA DE GADO DE CORTE (CAN)	187	10	CNPQ	342
11	EMBRAPA SEMIARIDO, CP 23, CEP 56.300-970 PETROLINA, PE.	183	11	EMBRAPA - SERVICO NACIONAL	329
12	EMBRAPA AMAZONIA OCIDENTAL,	156	12	CONSERVACAO DE SOLOS	311
13	1.2 EMBRAPA - CNPT, PASSO FUNDO, RS.	156	13	EMBRAPA - CNPA	284
14	CNPGL.	142	14	EMBRAPA - CPATU	263
15	EMBRAPA TRIGO, PASSO FUNDO, RS.	136	15	EMBRAPA - CCAA	243
16	EMBRAPA. CENTRO NACIONAL DE PESQUISA DE ARROZ E FEIJOAO (GO)	133	16	EMBRAPA - DEPARTAMENTO D	223
17	EMBRAPA.DID	116	17	EMBRAPA - CNPAF	218
18	EMBRAPA.DDT	105	18	EMBRAPA - CPPSE	210
19	1.2 EMBRAPA TRIGO, PASSO FUNDO, RS.	103	19	EMBRAPA - CPAFRO	208
20	EMBRAPA - UEPAE DE SAO CARLOS, SAO CARLOS, SP.	102	20	EMBRAPA - UEPAE DOURADOS	197
21	EMBRAPA PECUARIA SUDESTE, SAO CARLOS, SP.	100	21	DOCUMENTACAO	195
22	EMBRAPA. CENTRO NACIONAL DE PESQUISA DE GADO DE CORTE (CAN)	94	22	EMBRAPA - CNPMI	167

Fonte: Acervo Documental Embrapa, até 2000

Figure 6 : Production scientifique de l'Embrapa par centre (Penteado et al., 2002, p.12)

La création d'indicateurs de R&D de la recherche agronomique à partir de la base Collection documentaire de l'Embrapa, nécessitera une réforme des méthodes et routines de travail pour les adapter aux processus d'analyse automatique, comme l'a remarqué Van Raan (1997).

À partir du traitement bibliométrique de la base Collection documentaire, en addition aux indicateurs basiques comme nombre de travaux publiés par an, nombre de travaux publiés par unité de l'Embrapa et nombre de travaux publiés par chercheur, d'autres indicateurs peuvent être calculés, en substitution à des indicateurs du SAU/SAPRE comme :

Production Technique et Scientifique : 1.1-Article en Périodique Indexé ; 1.2-Chapître en Livre Technique et Scientifique ; 1.3-Article dans des Actes de Congrès / Note Technique ; 1.4-Résumé dans des Actes de Congrès ;

Production de Publications Techniques : 2.1-Circulaire Technique ; 2.2-Comunicat Technique ; 2.3-Bulletin de Recherche et Développement ; 2.4-Documents ; 2.5-Organisation/Édition de Livres ou de Système de Production.

Néanmoins, pour permettre l'utilisation de ces indicateurs par les systèmes de gestion, SISPAT, SAU et SAPRE, l'actualisation de la Collection documentaire, exécutée deux fois par an, au mois d'avril et novembre, quand sont ajoutées et réunies toutes les nouvelles insertions, devra être coordonnée avec le calendrier de la gestion pour qu'y soient inclus dans les bibliothèques, tous les travaux produits dans la période qui sera analysée. (...)

CONCLUSION

La bibliométrie est un outil important pour le processus de production d'informations pour la gestion des entreprises, en particulier, la génération d'indicateurs à partir de bases de données bibliographiques.

La méthodologie employée permet de travailler avec des données sûres et plus objectives, provenant d'une base bibliographique. Il permet aussi d'éviter la duplication d'efforts avec l'alimentation de différentes bases de données dans l'entreprise avec la même information. Il produit encore en plus de tableaux et de résultats graphiques qui permettent une meilleure visualisation des tendances.

Un autre résultat important obtenu avec ce travail est le nettoyage de la base Collection documentaire des inconsistances plus communes comme les entrées doublons d'auteurs et de titres.

En continuant ce travail dans la Collection documentaire, plusieurs autres indicateurs pourront être calculés et des analyses meilleures et plus profondes de la production scientifique de l'Embrapa sont à prévoir dans le futur, en intégrant d'autres sources d'information dans l'entreprise.

Le traitement de la collection documentaire et de la production scientifique

Comme l'ont indiqué Penteado et al. (2002), le travail bibliométrique offre deux produits spécifiques : D'abord, le nettoyage et le reformatage de la Collection documentaire de l'Embrapa (BDPA) suivant les recommandations de l'audit de 2002 et la norme ABNT NBR 6023 de 2000. Ensuite, le deuxième produit est l'analyse de la production scientifique et la génération de produits comme des indicateurs à partir de la base.

Pour accomplir ce travail, une étape importante a été l'achat du logiciel Infotrans, en juillet 2002 par l'Embrapa et l'entraînement du doctorant concomitant en son usage avec l'appui de l'Université Fédérale de Sao Carlos, NIT-Matériaux. Plusieurs améliorations ont été portées aux algorithmes et le processus d'indexation de données au logiciel Ainfo. En 2005, une autre étape fondamentale a été vaincue avec l'achat par l'Embrapa du logiciel VantagePoint. Ceci ouvre une capacité de diffusion du processus d'amélioration de la qualité et des audits des

bases bibliographiques en ce qui concerne toute l'entreprise avec le **VantagePoint Reader**, logiciel de libre distribution, avec le même concept de l'**Adobe Acrobat Reader**. Il permet aux centres de recherche d'accéder et d'explorer les analyses et audits de la base Collection documentaire et Production scientifique. Maintenant, nous décrivons le processus de nettoyage et reformatage de la BDPA.

Le QPAP (MPO, 2000), Pereira et al. (1999), Quoniam (1996) et Hood et Wilson (2000) ont confirmé que la question de la qualité des données est le point crucial de l'industrie de l'information. Rittberger et Rittberger (1997), Basch (1990), Wilson (1998), Medawar (1995), Fox et al. (1994), Armstrong (1994) et Jacsó (1997) soulignent que la construction de la qualité en bases de données concerne l'établissement de critères comme la couverture, l'accessibilité, l'actualisation, la consistance et la précision et que le développement accéléré de bases électroniques des ces dernières années n'a pas été toujours suivi ces critères. Cela étant, Blakeslee et Rumble Junior (2003) et Baer et al. (1995) indiquent que la meilleure solution est la planification et l'implémentation des processus qui assurent la présence de ces critères et principes de qualité dans les bases.

O'Neill et Vizine-Goetz (1988), le Gedoc (Sd), Medawar (1995), Mintz (1990), Dolan (1992), Casale (1993), Bourne (1977), Damerau (1964), Dalessandro (1986), Jacsó (1993a, 1993b, 1995, 1997), O'Neill et al. (1993), Pao (1989), Cahn (1994) et Jambu (2000) mentionnent que les erreurs les plus communs des bases de données concernent des problèmes d'orthographe, de dactylographie et d'omission. S'ajoutent aussi les erreurs de duplication focalisés par O'Neill et Vizine-Goetz (1988), Sule (1986), Williams et Maclaury (1979), Hickey et Rypka (1979), Mcpherson et al. (1982), Coyle et Gallaher-Brown (1985) et Ayres et al. (1984). Octaviano (2002) a rencontré que 94,5% des références de son échantillon de la Collection documentaire de l'Embrapa étaient des duplicata ou inconsistants. En conséquence, la première tâche à accomplir a été de travailler par le reformatage automatique d'Infotrans, dans la réduction des doublons et des inconsistances de la base.

Méthodologie

Les champs de la BDPA choisis pour ce travail sont décrits ensuite au tableau 7, suivis par un exemple du champ rempli.

Tableau 7 : Le registre Collection documentaire original

NR: Identificateur - Code d'identification du matériel - numération séquentielle - champ numérique = 51131

BI: Bibliothèque - abréviation - champ alphanumérique = CPAP

AU: Auteur - champ alphanumérique = BATISTA, M. de F.; PALAGI, P.M.

TI: Titre - champ alphanumérique = 'Bunchy top' da bananeira.

FO: Source - champ alphanumérique = Brasília: EMBRAPA-CENARGEN, 1985.

AN: Année de Publication - champ alphanumérique = 1985

OR: Origine du matériel - champ alphanumérique = EMB

MA: Type du matériel - champ alphanumérique = FL

PA: Mots clés - champ type mémo - max. 65.536 caractères = Banana; Bunch top; Doença; Vírus; Sintoma; Disease; Symptom

CO: Contenu - Décrit le contenu - résumé - abstrait - sommaire - champ type mémo - max. 65.536 caractères = Distribuição geográfica; Sintomas; Epidemiologia; Controle.

La base a été ensuite standardisée en enlevant tous les accents et symboles. Ils ont été substitués, par exemple, tous les Çç par Cc, Ññ par Nn, ÀÁÂÃÄàáâãä par A ou a et de même pour toutes les voyelles. On a constaté l'existence de références où l'un des champs étudiés n'était pas présent. Ainsi, un contrôle pour les champs vides a été établi comme le recommande Quoniam (1996), pour mesurer la qualité de chacun des champs étudiés, en déterminant le nombre de références renseignées (%). L'expression "**CampoVazio**" (ChampVide) a été introduite toujours quand le champ était vide ou non-existant dans le registre. Dans ce dernier cas, le champ était créé et l'expression introduite. La base Collection documentaire permet l'indexation de documents sans auteurs, seulement avec le titre. Ce-ci a permis d'obtenir un champs vide pour 8096 auteurs.

Il a été créé un nombre de registre unique pour chacun des 466.088 références. La liste de 268.035 auteurs de la base originale a été importée dans le logiciel VantagePoint, classée par fréquence et toutes les inconsistances ont été vérifiées jusqu'à la fréquence trois. Ensuite, la liste d'auteurs a été vérifiée par ordre alphabétique. Ce processus a été répété trois fois. Comme ont été identifiés au début 183.096 auteurs avec la fréquence un et 36.599 auteurs avec la fréquence deux, en additionnant les deux totales l'on trouve 219.694 auteurs, presque 82% de tous les auteurs, le critère de la proximité a été adopté quand l'analyse était faite par l'ordre alphabétique. Ces critères ont permis de déboucher sur une base contenant 152.021 auteurs à la fréquence un, soit une réduction de 17% et 32.271 auteurs à la fréquence deux,

soit une réduction de 11,8%. En somme, 80,4% des auteurs ont une fréquence comprise entre un et deux. Dans la liste de titres, 93,5% avaient une fréquence égale à un. En conséquence, la procédure adoptée a été celle de normaliser les noms des événements. Le nombre élevé de fréquences unitaires sera abordé dans le chapitre "Les limites du traitement".

Le programme de nettoyage de la BDPA comporte 45.728 lignes de code écrit. Les champs choisis comme prioritaires pour le nettoyage ont été les mêmes que ceux utilisés dans l'algorithme des duplicata : Auteur, titre et année de publication. Ces champs sont considérés comme très importants pour la bonne qualité de la base dès qu'ils déterminent la précision et le bon fonctionnement de l'algorithme.

Auteur

Dans le champ auteur, Octaviano (2002) a mentionné plusieurs inconsistances additionnelles. Entre parenthèses sont les solutions adoptées dans le nettoyage de la base avec le logiciel Infotrans : "&" entre auteurs et deux auteurs corporatifs comme entrée principale, séparés par "&" (solution = substituer "&" par ";"); Expression "ET AL." à la fin du champ (solution = chercher titre, année et source égales et inclure le champ avec les divers auteurs en substitution à l'original) ; Espace en blanc au début du champ et/ou à la fin (solution = retirer les espaces) ; Des auteurs orthographiés en minuscules (solution = passer pour majuscules) ; Auteurs avec entrée incorrecte de l'abréviation (solution = substituer l'indexation incorrecte par le nom correct de l'auteur) ; Local inséré après l'auteur corporatif comme en "SILICON GRAPHICS (Mountain View, CA, USA)." (solution = retirer le local de l'indexation, "SILICON GRAPHICS.") ; Trait d'union entre prénoms abrégés (solution = retirer le trait d'union) ; Éditeur, organisateur et d'autres, indiqués en minuscules et sans parenthèses (solution = substituer "ed.", "comp.", "org." et d'autres formes corrélâtes par "(Ed.)", "(Comp.)", "(Org.)") ; Désignatif "Junior" orthographié abrégé et en minuscules (solution = substituer JR et Jr par JUNIOR) ; Pays orthographié dans sa forme originale (solution = substituer le nom de Pays original, "UNITED STATES" par sa traduction en portugais, "ESTADOS UNIDOS") ; Divergences entre la façon d'orthographier un auteur, des auteurs non- identifiables et des auteurs cités de façon incomplète (solution = créer un champ "auteur corrigé" avec les corrections nécessaires) .

Pendant le développement du programme de reformatage de la BDPA, il a été décidé de faire un travail systématique de normalisation de noms d'institutions. D'abord, les centres de

recherche et départements de l'Embrapa, depuis les institutions brésiliennes, internationales, les événements nationaux et internationaux, les villes, états, pays et pour finir, les auteurs eux-mêmes.

Voici un échantillon de ce qui a été corrigé :

En noms d'institutions et inclus des centres de recherche de l'Embrapa: Erreurs d'orthographe, d'indexation et de dactylographie ;

Une institution internationale :

"IICA. Programa Cooperativo para el Desarrollo Tecnológico Agropecuario del Cono Sur."
(correct)

1. "IICA (Uruguay). Programa Cooperativo para el Desarrollo Tecnológico Agropecuario del Cono Sul."
2. "PROGRAMA COOPERATIVO PARA EL DESARROLLO TECNOLOGICO AGROPECUARIO DEL CONO SUR."
3. "PROGRAMA COOPERATIVO DE INVESTIGACION AGRICOLA DEL CONO SUR, IICA/BID/PROCISUR."
4. "PROGRAMA COOPERATIVO DE INVESTIGACION AGRICOLA DEL CONO SUR. IICA/BID/PROCISUR."
5. "PROCISUR. Montevideo, 1998."
6. "PROCISUR. Programa Cooperativo para el Desarrollo Tecnológico Agropecuario del cono sur."
7. "PROCISUR/IICA."
8. "Programa Cooperativo para el Desarrollo Tecnológico Agropecuario del cono sur."
9. "PROGRAMA COOPERATIVO DE INVESTACION AGRICOLA (Montevideo, Uruguay)."
10. "PROGRAMA COOPERATIVO DE INVESTIGACION AGRICOLA DE CONO SUR."
11. "PROGRAMA COOPERATIVO DE INVESTIGACION AGRICOLA DEL CONO SUR."
12. "PROGRAMA COOPERATIVO DE INVESTIGACION AGRICOLA. Convenio IICA Sur/BID."
13. "PROGRAMA COOPERATIVO DE INVESTIGACION AGRICOLA. Convenio IICA-Cono Sur/BID. Montevideo, Uruguay."
14. "Programa Cooperativo el Desarrollo Tecnológico Agropecuario del Cono Sur."
15. "PROGRAMA COOPERATIVO EL DESARROLO TECNOLOGICO AGROPECUARIO DEL CONO SUR."

16. "PROGRAMA COOPERATIVO PARA DESARROLLO TECNOLOGICO DEL CONO SUR."
17. "PROGRAMA COOPERATIVO PARA EL DESAGRO LLO TECNOLOGICO AGROPECUARIO DEL CONO SUR."
18. "PROGRAMA COOPERATIVO PARA EL DESAROLLO TECNOLOGICO AGROPECUARIO DEL CONO SUR."
19. "PROGRAMA COOPERATIVO PARA EL DESAROLLO TECNOLOGICO DEL CONO SUR."
20. "PROGRAMA COOPERATIVO PARA EL DESARROLLO AGROPECUARIO DEL CONO SUR."
21. "PROGRAMA COOPERATIVO PARA EL DESARROLLO TECNICO AGROPECUARIO DEL CONO SUR-PROCISUR."
22. "PROGRAMA COOPERATIVO PARA EL DESARROLLO TECNOLOGICO AGROPECUARIA DEL CONO SUL - PROCISUR."
23. "PROGRAMA COOPERATIVO PARA EL DESARROLLO TECNOLOGICO AGROPECUARIO DE CONO SUR. (PC 4768)."
24. "PROGRAMA COOPERATIVO PARA EL DESARROLLO TECNOLOGICO AGROPECUARIO DEL COMO SUR ARGENTINA-BOLIVIA-BRASIL-CHILE-PARAGUAY-URUGUAY."
25. "PROGRAMA COOPERATIVO PARA EL DESARROLLO TECNOLOGICO AGROPECUARIO DEL CONO SUL."
26. "Programa Cooperativo para el Desarrollo Tecnológico Agropecuario del Cono Sur."
27. "PROGRAMA cooperativo para el desarrollo tecnologico agropecuario del Cono Sur."
28. "PROGRAMA COOPERATIVO PARA EL DESARROLLO TECNOLOGICO AGROPECUARIO DEL CONOSUR."
29. "PROGRAMA COOPERATIVO PARA EL DESARROLLO TECNOLOGICO AGROPECUARIO."
30. "PROGRAMA COOPERATIVO PARA EL DESARROLLO TECNOLOGICO AGROPECUARIO DEL CONO SUR."
31. "PROGRAMA COOPERATIVO PARA EL DESARROLLO TECNOLOGICO DEL CONO SUR."
32. "PROGRAMA COOPERATIVO PARA EL DESARROLO TECNOLOGICO AGROPECUARIO DEL CONO SUR."
33. "PROGRAMA COOPERATIVO PARA EL DESARROLO TECNOLOGICO DEL CONO SUR."
34. "PROGRAMA COOPERATIVO PARA EL DESERROLLO TECNOLOGICO AGROPECUARIO DEL CONO SUR."
35. "PROGRAMA COOPERATIVO PARA EL DESERROLLO TECNOLOGICO AGROPECUARIO DEL CONO SUR."
36. "Programa Cooperativo para el Dessarrollo Tecnológico Agropecuario del Cono Sur."

37. "PROGRAMA COOPERATIVO PARA EL DESSAROLLO TECNOLOGICO AGROPECUARIO DEL CONO SUR."
38. "PROGRAMA COOPERATIVO PARA EL, DESARROLLO TECNOLOGICO AGROPECUARIA DEL CONO SUR."
39. "PROGRAMA COPERATIVO PARA EL DESARROLLO TECNOLOGICO AGROPECUARIO DEL CONO SUR."
40. "PROGRAMA COOPERATIVO DE INVESTIGACION AGRICOLA."
41. "PROCISUR-PROCISUR."
42. "PROCISUR PROCISUR."
43. "PROCISUR (PROCISUR)."

Une entreprise :

"SAS INSTITUTE." (correct)

1. "SAS INSTITUTE (Cary, NS.EUA)."
2. "SAS INSTITUTE INCORPORATION."
3. "SAS INSTITUTE INC."
4. "SAS Institute (Cary, NC)."
5. "SAS Institute Inc."
6. "SAS Institute."
7. "SAS Institut."
8. "SAS Institut, Cary-NC."
9. "SAS INSTITUTE INC. (Cary, N.C.)."
10. "SAS Institute (Caru, NC)."
11. "SAS Institute (Cary.)."
12. "SAS INSTITUTE (Cary, Estados Unidos)."
13. "SAS INSTITUTE (Cary, NC)."
14. "SAS Institute (Cary, NC, Estados Unidos)."
15. "SAS INSTITUTE INC., (Cary, North Carolina)."
16. "SAS INSTITUTE (Cary, North Carolina)."
17. "SAS INSTITUTE INC. (Cary, North Carolina)."
18. "SAS INSTITUTE INC (Cary, North Carolina)."
19. "SAS INSTITUTE (Cary, North Carolina, Estados Unidos)."
20. "SAS INSTITUTE (Cary, NS.EUA)."
21. "SAS Institute (Cary., NC)."
22. "SAS Institute Inc (Cary, NC)."

23. "SAS INSTITUTE INC. (Cary, CN)."
24. "SAS INSTITUTE INC. (Cary, Estados Unidos)."
25. "SAS INSTITUTE INC. (Cary, N.C.)."
26. "SAS INSTITUTE INC. (Raleigh, North Carolina)."
27. "SAS INSTITUTE INC.(Cary, NC)."
28. "SAS Institute Inc., North Carolina."
29. "SAS Institute(Cary, NC)."
30. "SAS INSTITUTE, (Cary, NC, Estados Unidos)."
31. "SAS Institute, (cory, N.C.)."
32. "SAS INSTITUTE, Cary, North Carolina."
33. "SAS. Institute Inc."
34. "SAS language and procedures: Introduction."
35. "SAS Language and Procedures: Syntax."
36. "SAS Language: reference."
37. "SAS Programming tips: a guide to efficient SAS Processing."
38. "SAS/GRAPH Software: Reference."
39. "SAS/STAT."

La normalisation des noms de divers événements comme congrès, séminaires, colloques, d'entre autres, nationales et internationales, concerne des entrées différentes pour un même événement :

Des erreurs d'orthographe et de dactylographie ;

2. Encontro sobre avaliacao socioeconomica da pesquisa **agopecuaria**.
2. Encontro sobre avaliacao socioeconomica da pesquisa **agropecauria**.
2. Encontro sobre avaliacao socioeconomica da pesquisa **agropecuaria**. (correct)

Des erreurs d'indexation ;

1. "ENCONTRO DA SOCIEDADE BRASILEIRA DE SISTEMAS DE PRODUCAO, 2., 1995." (correct).
2. "eNCONTRO DA SOCIEDADE BRASILEIRA DE SISTEMAS DE PRODUCAO."
3. "ENCONTRO DA SOCIEDADE BRASILEIRA DE SISTEMAS DE PRODUCAO, 1995,"
4. "Encontro da Sociedade Brasileira de Sistemas de Producao, 2., Londrina, PR, 1995"
5. "ENCONTRO DA SOCIEDADE BRASILEIRA DE SISTEMAS DE PRODUCAO, 1., 1996, Londrina."

6. "ENCONTRO DA SOCIEDADE BRASILEIRA DE SISTEMAS DE PRODUCAO, 1995, Londrina."
7. "ENCONTRO DA SOCIEDADE BRASILEIRA DE SISTEMAS DE PRODUCAO, 2., Londrina, 1995."
8. "Encontro da Sociedade Brasileira de Sistemas de Producao, 2., Londrina, PR, 1995"

Finalment, un échantillon du travail avec les noms d'auteurs : Erreurs d'orthographe, d'indexation et de dactylographie ;

"HERRMANN JUNIOR, P. S. de P.;" (correct)

1. "HERMANN JUNIOR, P. S.;"
2. "HERMANN JUNIOR, P. S. de P.;"
3. "HERMANN, P."
4. "HERMANN, P. S. P.;"
5. "HERRMANN, P."
6. "HERRMANN, P. S.;"
7. "HERRMANN, P. S. P.;"
8. "HERRMANN, P. S. de P.;"
9. "HERRMANN JUNIOR, P. S. P."
10. "HERRMANN JUNIOR, S. de P."

Voici un autre exemple :

"REIFSCHNEIDER" (correct)

"REIFSCHNEIDER, F. J. B." (correct)

1. "REIFDCHNEIDER"
2. "REIFSCHENEIDER"
3. "REIFSCHINEIDER"
4. "REIFSCHNEIRDER"
5. "REIFSCHNEIDER; F. J. B.;"
6. "REIFSCHNEIDER, F."
7. "REIFSCHNEIDER, F. B."
8. "REIFSCHNEIDER, F. J."
9. "REIFSCHNEIDER, F. J. S."
10. "REIFSCHNEIDER, J. B."
11. "REIFSCHNEIDER, J. B. F."
12. "REIFSCHSNEIDER, F."

13. "REIFSCHSNEIDER, F. J. B."
14. "REIFSHNEIDER, E. F. R."
15. "REIFSHNEIDER, F. J. B."
16. "REISFCHNEIDER, F. J. B."
17. "REISFSCHNEIDER, F. J. B."

Titre

Dans le champ "titre" les inconsistances mentionnées par l'audit (et les solutions adoptées) sont : Catalogage par le titre en minuscules, titre tout en majuscules et mot à la fin du titre en majuscule (solution = transformer les deux premiers mots du titre en majuscules et les autres en minuscules) ; Titre entre guillemets (solution = retirer les guillemets) ; Absence de titre, titre comme catalogage principal et répété dans le champ titre (solution = identifier registre pour correction manuelle).

La majorité absolue des erreurs sont d'orthographe, d'indexation et de dactylographie. Quelques entrées, comme les onze ci-dessus, bien qu'étant apparemment différentes, reviennent comme égales pour l'algorithme d'identification de doublons parce qu'il ignore les signaux graphiques, espaces et lettres majuscules ou minuscules.

1. "Arquivistica = tecnica: Arquivologia = ciencia."
2. "Arquivista = tecnica: arquivologia = ciencia."
3. "Arquivistica = tecnica arquivologia = ciencia."
4. "Arquivistica = tecnica: arquivologia = ciência."
5. "Arquivistica tecnica: arquivologia-ciencia."
6. "Arquivistica- Tecnica: Arquivologia- Ciencia."
7. "Arquivistica: tecnica, arquivologia ciencia."
8. "Arquivistica: Tecnica: arquivologia: ciencia."
9. "Arquivistica= tecnica: arquivologia-ciencia."
10. "Arquivistica=Tecnica, Arquivologia=Ciencia."
11. "Arquivistica=tecnica: aquivologia=ciencia."

Par contre, dans les exemples suivants, même s'ils sont des références d'un même document, l'algorithme des doublons ne fonctionne pas avec toute son efficacité :

1. "Anais do V Congresso Brasileiro de Ciencia do Solo."
2. "Anais do 5. Congresso Brasileiro de ciencia do solo."
3. "Anais do quinto congresso brasileiro de ciencia do solo."

Voici un autre exemple :

1. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13., 1991; SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO, 1., 1991; ENCONTRO SOBRE COCHONILHA DA PALMA FORRAGEIRA, 2., 1991; ENCONTRO SOBRE MOSCAS-DAS-FRUTAS, 3., 1991." (correct)
2. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13 & I SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO II ENCONTRO SOBRE COCHONILHA DA PALMA FORRAGEIRA III ENCONTRO SOBRE MOSCAS-DAS-FRUTAS. Recife, 1991."
3. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13, & I SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO II ENCONTRO SOBRE COCHONILHA DA PALMA FORRAGEIRA III ENCONTRO SOBRE MOSCAS-DAS-FRUTAS. Recife, Recife."
4. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13., Recife, 1991. e SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO, 1. e ENCONTRO SOBRE COCHONILHA DA PALMA FORRAGEIRA, 2. e ENCONTRO SOBRE MOSCAS DAS FRUTAS, 3."
5. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13., SIMOPOSIO INTERNACIONAL SOBRE "BICUDO" DO ALGODOEIRO, 1., 1991. Recife."
6. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13., SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO, 1., 1991. Recife."
7. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13.; SIMPOSIO INTERNACIONAL SOBRE "BICUDO" DO ALGODOEIRO, 1.; ENCONTRO SOBRE "COCHONILHA" DA PAIMA FORRAGEIRA, 2.; ENCONTRO SOBRE MOSCA-DAS-FRUTAS, 3., 1991, Recife."
8. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13.; SIMPOSIO INTERNACIONAL SOBRE 'BICUDO' DO ALGODOEIRO, 1.; ENCONTRO SOBRE 'COCHONILHA' DA PALMA FORRAGEIRA, 2.; ENCONTRO SOBRE MOSCAS-DAS-FRUTAS, 3., 1991, Recife."
9. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13.; SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO, 1.; ENCONTRO SOBRE COCHONILHA DA PALMA FORRAGEIRA, 2.; ENCONTRO SOBRE MOSCAS DAS FRUTAS, 3., 1991, Recife."
10. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13.; SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO, 1.; ENCONTRO SOBRE COCHONILHA DA PALMA FORRAGEIRA, 2.; ENCONTRO SOBRE MOSCAS - DAS - FRUTAS, 3., 1991, Recife, PE."
11. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13.; SIMPOSIO INTERNACIONAL SOBRE "BICUDO" DO ALGODOEIRO, 1.; ENCONTRO SOBRE "COCHONILHA" DA PALMA FORRAGEIRA, 2.; ENCONTRO SOBRE MOSCAS-DAS-FRUTAS, 3., 1991, Recife."
12. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13.; SIMPOSIO INTERNACIONAL SOBRE "BICUDO" DO ALGODEIRO, 1.; ENCONTRO SOBRE

- "COCHONILHA" DA PALMA FORRAGEIRA, 2; ENCONTRO SOBRE MOSCA-DAS-FRUTAS, 3., 1991, Recife."
13. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13.; SIMPOSIO INTERNACIONAL SOBRE "BICUDO" DO ALGODOEIRO, 1.; ENCONTRO SOBRE "COCHONILHA" DA PALMA FORRAGEIRA, 2.; ENCONTRO SOBRE MOSCA-DAS-FRUTAS, 3., 1991, Recife."
 14. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13 & I SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO II ENCONTRO SOBRE COCHONILHA DA PALMA FORRAGEIRA III ENCONTRO SOBRE MOSCAS-DAS-FRUTAS. Recife, 1991."
 15. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13, & I SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO II ENCONTRO SOBRE COCHONILHA DA PALMA FORRAGEIRA III ENCONTRO SOBRE MOSCAS-DAS-FRUTAS. Recife, Recife."
 16. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13, 1991, Recife."
 17. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13."
 18. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13., Recife, 1991. e SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO, 1. e ENCONTRO SOBRE COCHONILHA DA PALMA FORRAGEIRA, 2. e ENCONTRO SOBRE MOSCAS DAS FRUTAS, 3."
 19. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13., SIMOPOSIO INTERNACIONAL SOBRE "BICUDO" DO ALGODOEIRO, 1., 1991. Recife."
 20. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13., SIMPOSIO INTERNACIONAL SOBRE BICUDO DO ALGODOEIRO, 1., 1991. Recife."
 21. "CONGRESSO BRASILEIRO DE ENTOMOLOGIA, 13. Recife, PE. 1991."

La normalisation des noms d'événements a été le principal travail pour conférer plus d'efficience à l'algorithme d'identification des doublons.

Année de publication

Ce champ est particulièrement important pour l'algorithme des doublons parce qu'il est le premier item d'identification de similarités et différences. Les principaux problèmes sont aussi des erreurs d'orthographe, d'indexation, de dactylographie et les champs vides. Quand l'année n'est pas précisée dans la fiche de catalogage des publications, la norme ancienne préconisait de laisser le champ vide. Ceci est en contradiction avec la doctrine sur les erreurs d'omission (Medawar, 1995 ; Jacsó, 1993a et 1993b ; Mintz, 1990).

Ainsi, dans la base originale il y avait 36.253 références avec l'année en champ vide, 2320 avec le nombre 0, 100 avec le nombre 19, 50 références ayant comme année de publication le nombre 198, 46 avec le 197, 11 avec le 196, 20 références avec les chiffres 94, 1, 200, 93 et

195 et autres 114 avec les fréquences un ou deux qui ont été considérés comme douteux. Comme le champ en question était numérique mais n'était pas choisi à partir d'une liste de chiffres d'années, il ouvrait la possibilité d'introduire d'erreurs au moment de l'indexation.

Ces 38.920 références considérées comme douteuses représentaient 8,3% du total. En les analysant nous avons constaté que beaucoup d'informations correspondant à l'année de publication étaient présentes de façon complète dans le champ "source". Ainsi, le programme de nettoyage cherchait l'information concernant l'année de publication dans le champ "source" et le comparait à l'information du champ lui-même. Quand le champ était vide ou avec une information non fiable comme "2988", "1)84" et "9889", l'information du champ "source" était introduite. Il a été considéré aussi qu'un bon critère pour l'approximation était celui de la décade de la publication qui est plus proche que le siècle de publication. De la même façon, il y avait deux dénominations pour l'information non existante: ChampVide et le nombre zéro (0). Ce dernier a été adopté pour indiquer l'absence d'information sur l'année.

Pour Quoniam (1996) un des critères de qualité d'une base de données est la quantité de références sans informations, les champs vides. Ces critères ont donné plus de consistance aux résultats. Dans la base originale 227 formes étaient vides. Dans la base traitée il y a eu une réduction de 25%, pour 171 formes. La forme 0 (zéro) qui, dans la base originale était présente dans 2.320 références a augmenté pour concerner 4.922 références mais 33.998 références considérées douteuses dans la base originale se sont distribués de manière presque uniforme entre les autres années. Les formes avec fréquence 1 et 2 se sont réduites de 114 pour 45. Les 45 formes les plus fréquentes sont présentés tableau 8. Les indications de champs vides sont marquées en gris. La colonne "# Références" indique la fréquence de la forme ou, selon Rostaing (SD, p.12), "le nombre de références où la forme est présente". L'expression "# Occurrences" indique la "fréquence des occurrences de la forme, c'est-à-dire, le nombre de fois que la forme apparaît dans l'ensemble des références" (Rostaing, Sd., p.12).

Tableau 8 : Résultat du reformatage du champ "année de publication"

Base Collection Originale			Base Collection Traitée		
année de publication			année de publication		
# Occurrences	ANO:		# Occurrences	ANO:	
1	36253	ChampVide	1	12896	2000
2	12378	2000	2	11831	1998
3	11360	1998	3	11684	1982
4	11071	1982	4	11495	1997
5	11056	1997	5	11367	1999
6	10953	1999	6	11068	1980
7	10534	1996	7	10920	1996
8	10479	1995	8	10712	1983
9	10151	1981	9	10677	1995
10	10136	1980	10	10662	1984
11	10068	1984	11	10621	1981
12	10061	1994	12	10615	1992
13	10053	1983	13	10401	1994
14	9839	1992	14	10222	1993
15	9434	2001	15	9823	1986
16	9347	1993	16	9788	2001
17	9121	1979	17	9774	1979
18	9051	1978	18	9687	1978
19	8736	1988	19	9454	1990
20	8692	1986	20	9364	1988
21	8690	1991	21	9246	1991
22	8538	1989	22	8971	1985
23	8274	1985	23	8852	1989
24	8273	1990	24	8462	2002
25	8118	2002	25	8405	1987
26	7550	1977	26	8115	1977
27	7504	1987	27	7322	1976
28	6767	1976	28	6388	1975
29	5865	1975	29	5235	1974
30	4576	1974	30	4922	0
31	4146	1973	31	4526	1973
32	3791	1972	32	4104	1972
33	3561	2003	33	3762	1971
34	3511	1971	34	3657	1970
35	2898	1970	35	3544	2003
36	2370	1969	36	3037	1900
37	2320	0	37	2669	1969
38	2114	1968	38	2338	1968
39	1832	1967	39	2050	1967
40	1798	1966	40	2035	1966
41	1356	1965	41	1596	1965
42	1262	1964	42	1418	1964
43	1129	1961	43	1306	1963
44	1128	1963	44	1228	1961
45	1078	1960	45	1224	1960

Doublons

En concernant les doublons, des références qui n'ont pas été identifiées par l'algorithme, Octaviano souligne qu' "ils ont leur origine dans des erreurs d'indexation, de dactylographie et de remplissage des champs" et qu'à de nombreuses reprises le nombre de doublons non identifiés par l'algorithme "a dépassé une dizaine" (2002, p.8). Selon l'auditrice, "on remarque aussi le placement de données dans d'autres champs qui ne sont pas propres" et "des catalogages divergents pour un même document, (...) entré par le titre, par l'auteur principal et par l'auteur corporatif, ou avec un même auteur de trois façons différentes" (2002, p.78).

Le travail de nettoyage a été réalisé, comme mentionné, en identifiant et en corrigeant des fautes qui réduisaient l'efficacité de l'algorithme des doublons. Ainsi, pour évaluer l'efficacité d'un nouvel algorithme d'identification de doublons, il a été décidé de travailler avec la base entière, c'est-à-dire, toutes les références valides de la BDPA, dans un total de 466.088 références, situation de novembre 2003. La base de travail a été construite de manière à profiter de l'algorithme d'élimination de doublons d'Infotrans. Cet algorithme préserve toujours le dernier registre trouvé dans la base parmi les duplicata. En suivant les indications de l'audit et les observations de travail de la CP-Ainfo, les centres de recherche ont été classés selon leur qualité d'indexation des références. La base a été alors, remontée en additionnant chaque collection de centre de recherche, en première place les collections avec la moindre qualité d'indexation et en dernière place celles avec la meilleure qualité d'indexation. De la même manière, à la fin de la base ont été séparés et classés les références indiquées par les centres de recherche dans le champ "OR" - "origine" - comme faisant partie de leur "production scientifique".

Pour différencier entre des matériaux de types divers, comme des livres et des CÉDÉROM, par exemple, mais avec le même auteur et titre, le contenu du champ "MA" - "type de matériel" - a été inclus dans le code unique de contrôle du registre, ici appelé **dubcode**. Le **dubcode** a été construit en sélectionnant les deux premiers caractères des 18 premières formes du champ "auteur" (1-18AU) et les quatre derniers caractères de ce même champ (4DAU), les deux premiers caractères des 21 premières formes du champ "titre" (1-21TI) et les quatre derniers caractères de ce même champ (4DTI), les deux derniers caractères du champ "année de publication" (2AN) et les deux caractères du champ "type de matériel" (2MA). Au total le **dubcode** a 38 caractères qui ont été montés d'une façon non conventionnelle, en inversant

l'ordre des formes du champ "auteur" et du champ "titre" et en enlevant tous les espaces en blanc à la fin du processus. La formule du **dubcode** est la suivante:

$$\begin{aligned} \text{Dubcode} = & (2\text{AN}) + (2\text{MA}) + (21\text{TI}) + (12\text{TI}) + (13\text{TI}) + (3\text{TI}) + (4\text{DTI}) + (17\text{TI}) + \\ & (16\text{TI}) + (14\text{TI}) + (11\text{TI}) + (10\text{TI}) + (9\text{TI}) + (8\text{TI}) + (7\text{TI}) + (6\text{TI}) + (5\text{TI}) + (4\text{TI}) + (2\text{TI}) + \\ & (1\text{TI}) + (18\text{AU}) + (16\text{AU}) + (10\text{AU}) + (8\text{AU}) + (7\text{AU}) + (6\text{AU}) + (5\text{AU}) + (4\text{AU}) + (3\text{AU}) + \\ & (2\text{AU}) + (1\text{AU}) + (4\text{DAU}) \end{aligned}$$

Le chapitre 5 du manuel de l'Infotrans (IUK, 1995, pp.3-4) cite un exemple de construction de **dubcode** en sélectionnant les caractères de chaque forme et en les disposant dans leur ordre naturel de disposition, pour les peuples issus des civilisations greco-romaines, de la gauche à la droite. Comme il existait à la BDPA des références avec des titres très longs et des multiples auteurs, il s'est avéré nécessaire de sélectionner, au moins 21 formes du champ "titre" et autres 18 formes dans le champ "auteur". Alors dans plusieurs cas, les éléments qui différenciaient auteurs et titres étaient dans les dernières façons. En conséquence, l'option a été faite par l'inversion de l'ordre naturelle et de mettre les derniers éléments en première place.

Ainsi des **dubcodes** qui en suivent l'ordre naturel, auraient une longue séquence de caractères identiques et seulement des lettres changées à la fin, identifient des références comme différents déjà au début de l'algorithme, ce qui rend ce processus de comparaison plus performant. Notre conviction c'est de que ce principe doit être utilisé toujours quand on observe des champs très longs.

Exécution

Le programme de nettoyage de la BDPA a compté huit archives pour le traitement du champ "auteur" (aut0, aut1, aut1a, aut2, aut3, aut4, aut5, aut6) et deux archives (tit0, tit1) pour traiter les autres champs : Titre, source et année de publication. La base traitée était passée alors par le programme (*gdub) générateur des **dubcodes**. Ensuite, les **dubcodes** étaient comparés par l'Infotrans. Les doublons identifiés étaient séparés dans un archive avec leurs chiffres uniques d'identification et leurs bibliothèques d'origine (affiliation). Ces deux informations étaient insérées en la base à la fin du processus.

Le programme de nettoyage écrit pour le logiciel Infotrans a été tourné dans un microprocesseur Atlon avec **clock** de 1.7 Mégahertz, 384 Mégabits de mémoire RAM et

système opérationnel Windows 98-SE. Dans cette configuration le temps d'exécution de toute la procédure a été d'environ 120 heures ou cinq jours d'opération.

À la fin du processus, la base de données Collection documentaire (BDPA) avait les champs suivants, suivis par un exemple du champ rempli, en supposant que deux duplications seraient trouvées. Voyez dans le tableau 9.

Tableau 9 : Le registre Collection documentaire traité

REG: Identificateur - champ numérique = 51131
NRG: Nombre unique du registre - champ numérique = 1, 47700, 290876
AUO: Auteur original - champ alphanumérique = BATISTA, M. de F.; PALAGI, P.M.
AUT: Auteur traité - champ alphanumérique = BATISTA, M. de F.; PALAGI, P. M.
TIO: Titre original - champ alphanumérique = 'Bunchy top' da bananeira.
TIT: Titre traité - champ alphanumérique = Bunchy top da bananeira.
PUB: Source - champ alphanumérique = Brasilia: EMBRAPA-CENARGEN, 1985.
JOR: Revue de Publication - champ alphanumérique = CampoVazio
AN0: Année de Publication originale - champ alphanumérique = 1985
ANO: Année de Publication traitée - champ alphanumérique = 1985
UDS: Bibliothèque - abréviation - champ alphanumérique = CPAP, CNPGC, CENARGEN
PRO: Origine du matériel - champ alphanumérique = UPC
MAT: Type du matériel - champ alphanumérique = FL
PCH: Mots clés - champ type mémo - max. 65.536 caractères = Banana; Bunchy top; Doenca; Virus; Sintoma; Disease; Symptom
CON: Contenu - Décrit le contenu - résumé - abstrait - sommaire - champ type mémo - max. 65.536 caractères = Distribuicao geografica; Sintomas; Epidemiologia; Controle.

Résultats du nettoyage

Le programme de nettoyage de la BDPA a trouvé 113.166 références doublons, 24,28% des 466.088 références. Il a trouvé 9,4% plus de références doublons que l'algorithme originel.

Un total de 219.694 auteurs avait la fréquence un ou deux dans la base originale, 183.096 auteurs avec la fréquence "un" et 36.599 avec la fréquence "deux". Ces listes ont été mises en ordre alphabétique et le critère de proximité a été utilisé pour l'identification des inconsistances. De ces critères ont résulté, dans la base traitée, 152.021 auteurs avec la fréquence "un", soit une réduction de 17% et 32.271 auteurs avec la fréquence "deux", soit

une réduction de 11,8%. Les 228 années de publication de la base originale ont été réduites à 171. De la même forme, 308.299 titres se sont réduits à 304.881 et 268.035 auteurs à 228.976. Ce dernier nombre représente moins 39.059 auteurs (14,78%) et a indiqué des altérations significatives dans les statistiques de comptage du nombre de publications par auteur. Le tableau 10 présente l'exemple des 56 premiers auteurs et les 56 premiers de la base traitée, en accompagnant les chiffres de la dernière colonne à droite qui indique la classification de l'auteur dans la base originale.

Tableau 10 : Auteurs, base Collection documentaire originale et traitée

Base Collection originale			Base Collection traitée			
# Occurrences	Auteur		# Occurrences	Auteur	ClassOri	
1	8096	ChampVide	1	8096	ChampVide	1
2	1139	EMBRAPA (Brasilia, DF	2	1768	EMBRAPA	2
3	650	DOBEREINER, J	3	1035	FAO	4
4	529	FAO (Roma, Italia	4	728	KIIHL, R. A. de S	6
5	478	CRUZ, I	5	691	ALMEIDA, L. A. de	12
6	468	KIIHL, R.A.S	6	595	EMBRATER	14
7	423	EMBRAPA. Cerrados (Planaltina, DF	7	586	DOBEREINER, J	3
8	411	CARDOSO, M.J	8	542	SANTOS, M. X. dos	30
9	409	YORINORI, J.T	9	510	YORINORI, J. T	9
10	408	LOPES, C.A	10	499	FRANCA NETO, J. de B	19
11	398	REIFSCHNEIDER, F.J.B	11	474	HENNING, A. A	18
12	392	ALMEIDA, L.A	12	472	GAMA, E. E. G	39
13	389	COSTA, N. de L	13	471	EMBRAPA. Cerrados	7
14	386	EMBRATER (Brasilia, DF	14	459	CRUZ, I	5
15	375	GARCIA, A	15	454	LOPES, C. A	10
16	354	HUNGRIA, M	16	451	REIFSCHNEIDER, F. J. B	11
17	339	MOSCARDI, F	17	442	BELTRAO, N. E. de M	23
18	338	HENNING, A.A	18	426	COSTA, N. P. da	94
19	331	FRANCA NETO, J. de B	19	420	GAZZIERO, D. L. P	24
20	328	EMBRAPA	20	418	CARDOSO, M. J	8
21	325	MALAVOLTA, E	21	395	ALENCAR, M. M. de	27
22	325	PARENTONI, S.N	22	390	OLIVEIRA, A. C. de	67
23	324	BELTRAO, N.E. de M	23	388	COSTA, N. de L	13
24	321	GAZZIERO, D.L.P	24	381	GIORDANO, L. de B	60
25	316	CRESTANA, S	25	380	IBGE	61
26	316	PRIMAVESI, O	26	374	KRZYZANOWSKI, F. C	40
27	315	ALENCAR, M.M. de	27	373	GARCIA, A	15
28	305	SIMPLICIO, A.A	28	367	SFREDO, G. J	29
29	299	SFREDO, G.J	29	355	BORKERT, C. M	32
30	296	SANTOS, M.X. dos	30	340	CARVALHO, H. W. L. de	31
31	292	CARVALHO, H.W.L. de	31	340	MOSCARDI, F	17
32	291	BORKERT, C.M	32	339	HUNGRIA, M	16
33	284	PACHECO, C.A.P	33	330	FREITAS, A. R. de	59
34	283	FRANCO, A.A	34	326	SUDENE	350
35	272	FREIRE, E.C	35	319	CRESTANA, S. A	25
36	272	VASCONCELLOS, C.A	36	314	LANTMANN, A. F	37
37	266	LANTMANN, A.F	37	314	SIMPLICIO, A. A	28
38	261	VOLL, E	38	310	PRIMAVESI, O. M. A. S. P. R. B	26
39	257	GAMA, E.E.G	39	308	MALAVOLTA, E	21
40	252	KRZYZANOWSKI, F.C	40	308	PARENTONI, S. N	22
41	252	MAGALHAES, P.C	41	308	RIBEIRO, V. Q	45
42	251	CRUZ, J.C	42	307	CARVALHO, P. E. R	54
43	249	MAGNAVACA, R	43	299	FRANCO, A. A	34
44	248	PAIVA, E	44	298	SILVA, J. B. da	297
45	243	RIBEIRO, V.Q	45	297	AVILA, A. C. de	113
46	238	GASPAROTTO, L	46	294	FREIRE, E. C	35
47	238	GODOY, R	47	292	AZEVEDO, D. M. P. de	62
48	238	TORRES, E	48	292	EMBRAPA. Solos	55
49	237	FRANCA, F.H	49	290	JUNQUEIRA, N. T. V	87
50	235	CAMPO, R.J	50	281	ALBUQUERQUE, F. C. de	84
51	235	CASALI, V.W.D	51	281	SOUZA, P. I. M	342
52	234	EUCLIDES FILHO, K	52	277	CAMPO, R. J	50
53	232	MALUF, W.R	53	274	VASCONCELOS, C. A	36
54	229	CARVALHO, P. E. R	54	269	PACHECO, C. A. P	33
55	227	EMBRAPA. Solos (Rio de Janeiro, RJ	55	269	PRIMAVESI, A. C. P. de A	248
56	227	WAQUIL, J.M	56	268	IICA	598

Note. Class Ori. = Classification base originale.

Le tableau 11 indique les 20 auteurs qui ont le "plus bénéficiés" du reformatage, nettoyage et normalisation des noms. Entre eux, trois institutions, le IICA, la SUDENE et le SAS INSTITUTE qui respectivement étaient dans les 598°, 350° et 162° positions et ont passé à 55°, 33° et 80° positions respectivement. Du coté des auteurs personnes physiques, FRANCA, G. E. de; CARVALHO, M. M. de; SOUZA, P. I. M. et SILVA, J. B. ont passé respectivement de 429°, 354°, 342° et 297° places pour 88°, 83°, 49° et 43° respectivement.

Tableau 11 : Les 20 auteurs le plus "bénéficiés" par le reformatage

Base Collection Traitée				
	#	Auteur	Class	Class
	Occurrences		Orig.	Traitée
1	268	IICA	598	55
2	222	FRANCA, G. E. de	429	88
3	225	CARVALHO, M. M. de	354	83
4	326	SUDENE	350	33
5	281	SOUZA, P. I. M	342	49
6	298	SILVA, J. B. da	297	43
7	269	PRIMAVESI, A. C. P. de A	248	53
8	220	SILVA, W. L. C	224	93
9	245	CAMPOS, O. F. de	220	67
10	221	GUIMARAES, P. E. O	212	90
11	233	SAS INSTITUTE	162	80
12	221	MARQUELLI, W. A	149	91
13	220	RESENDE, G. M. de	146	94
14	211	BOITEUX, L. S	141	99
15	219	ALMEIDA, D. L. de	139	95
16	215	SILVA, J. F. V	134	97
17	212	HAJI, F. N. P	127	98
18	224	NASCIMENTO, W. M	124	85
19	297	AVILA, A. C. de	113	44
20	235	ASSAD, E. D	112	79

Note. Class Orig. = Classification base originale.
Class Traitée = Classification base traitée.

Limites du travail

Les limites du traitement automatique

La Collection documentaire continue avec un grand nombre de fréquences unitaires, 80,5% en auteur et 93,5% en titres. Une rapide évaluation de ces listes comme dans le tableau 12, montre avec clarté qu'il y a encore beaucoup de travail à faire sur la BDPA. Nous nous intéressons de façon privilégiée aux champs figurant en gras dans le tableau 12.

Tableau 12 : Petite échantillon des 152.021 auteurs et 284.950 titres avec fréquence un

Auteurs Fréquence=1	Titres Fréquence=1
ABREU, A	Abundancia, frecuencia e grau de agregacao do pau rosa Aniba Duckei Kostermans na floresta de Tapajos
ABREU, A . A. de. (Coord	Abundancia, frecuencia e grau de agregacao do pau-rosa (Aniba duckei Kostermans) na floresta nacional do Tapajos
ABREU, A. G	Abundancia, frecuencia e grau de agregacao do pau-rosa (Aniba duckei) na Floresta Nacional do Tapajos
ABREU, A. V. de	Abundante a frutificacao da Imbuia
ABREU, C. A	Ac conductivity behavior of poly (o-methoxyaniline
ABREU, C. A. (Ed)	Ac conductivity of poly (o-methoxyaniline
ABREU, C. A. de. (Ed)	Ac-70-154, a gummy stem blight resistant muskmelon breeding line with desirable horticultural characteristics
ABREU, C. B	Ac-70-154, a gummy stem blight resistant muskmelon breeding line with desirable horticultural characteristics
ABREU, C. C. de	Ac-70-154, a gummy stem blight-resistant muskmelon breeding line
ABREU, C. L. M	Acabamento de caprinos durante a epoca seca no Nordeste do Brasil
ABREU, D. de	Acabamento de caprinos e ovinos durante a epoca seca, no Nordeste

Dans la liste Auteurs à fréquence un, à gauche du tableau 12, nous avons remarqué les auteurs ABREU, C. A.; ABREU, C. A. (Ed) et ABREU, C. A. de. (Ed.). Les références sont présentées tableau 13.

Tableau 13 : Trois références similaires (ABREU, C. A.)

ANO: 1999
AUT: FURLANI, P. R. ; FURLANI, A. M. C. ; ABREU, M. F. ; ABREU, C. A

PUB: In: ENCONTRO NACIONAL SOBRE SUBSTRATO DE PLANTAS, 1., 1999, Porto Alegre. Substrato para plantas: a base da producao em recipientes: anais... Porto Alegre: Genesis, 1999

REG: 336391

TIT: Caracterizacao quimica de substrato e o desenvolvimento de mudas de tomateiro

UDS: CNPH

ANO:2001

AUT:RAIJ, B. van. ; FERREIRA, M. E. ; CRUZ, M. C. P. da. ; ABREU, C. A. (Ed)

PUB: Joticabal: CNPq / FAPESP / POTAFOS, 2001

REG:343623

TIT: Micronutrientes e elementos toxicos na agricultura

UDS: CNPH

ANO: 2001

AUT: RAIJ, B. van. ; FERREIRA, M. E. ; CRUZ, M. C. P. da. ; ABREU, C. A. de. (Ed)

PUB: Jaboticabal: CNPq / FAPESP / POTAFOS, 2001

REG: 374429

TIT: Micronutrientes e elementos toxicos na agricultura

UDS: CNPSO

Après analyse, nous déduisons que les deux derniers sont doublons.

Dans la liste des Titres à fréquence un, à droite du tableau 12, trois sont très similaires : Abundancia, frecuencia e grau de agregacao do pau rosa Aniba Duckei Kostermans na floresta de Tapajós; Abundancia, frecuencia e grau de agregacao do pau-rosa (Aniba duckei Kostermans) na floresta nacional do Tapajós e Abundancia, frecuencia e grau de agregacao do pau-rosa (Aniba duckei) na Floresta Nacional do Tapajós. Regardez-les trois références au tableau 14.

Tableau 14 : Trois titres similaires

ANO: 1983

AUT: CARVALHO, J. O. P. de

PUB: Belem: EMBRAPA-CPATU, 1983

REG: 165892

TIT: Abundancia, frecuencia e grau de agregacao do pau rosa Aniba Duckei Kostermans na floresta de Tapajos

UDS: CENARGEN

ANO: 1983

AUT: CARVALHO, J. O. P. de

PUB: Belem: EMBRAPA-CPATU, 1983

REG: 436956

TIT: Abundancia, frecuencia e grau de agregacao do pau-rosa (Aniba duckei Kostermans) na floresta nacional do Tapajos

UDS: CPATU

ANO: 1983

AUT: CARVALHO, J. O. P. de

PUB: Belem: EMBRAPA-CPATU, 1983

REG: 90072

TIT: Abundancia, frecuencia e grau de agregacao do pau-rosa (Aniba duckei) na Floresta Nacional do Tapajos

UDS: CPAMN

Nous avons conclu que ce sont des doublons.

Au tableau 12 nous avons aussi remarqué des similarités dans trois d'autres titres : Ac-70-154, a gummy stem blight-resistant muskmelon breeding line; Ac-70-154, a gummy stem blight resistant muskmelon breeding line with desirable horticultural characteristics e Ac-70-154, a gummy stem blight resistant muskmelon breeding line with desirable horticultural characteristic. Ils sont présentés tableau 15.

Tableau 15 : Autres trois titres similaires du tableau 12

ANO: 1989

AUT: NORTON, J. D. ; COSPER, R. D

PUB: HortScience, v. 24, n. 4, p. 709-711, Aug., 1989

REG: 90022

TIT: Ac-70-154, a gummy stem blight-resistant muskmelon breeding line

UDS: CPAMN

ANO: 1985
 AUT: NORTON, J. D. ; COSPER, R. D
 PUB: Cucurbit Genetics Cooperative Report, n. 8, p. 46-49, 1985
 REG: 90021
 TIT: Ac-70-154, a gummy stem blight resistant muskmelon breeding line with desirable horticultural characteristics
 UDS: CPAMN

ANO: 1985
 AUT: NORTON, J. D.; COSPER, R. D
 PUB: Report. Cucurbit. Genetics Cooperative, n. 8, p. 46-49, 1985
 REG: 79925
 TIT: Ac-70-154, a gummy stem blight resistant muskmelon breeding line with desirable horticultural characteristic
 UDS: CPAMN-UEPP

Nous déduisons que le premier est différent des autres deux qui sont des doublons.

Le programme de nettoyage n'a pas été capable de retirer tous les doublons. En réalité, le travail doit et nécessite d'être poursuivi au niveau des bibliothèques des centres de recherche. Ceci parce qu'une bonne partie des problèmes qui restent ont besoin du jugement et de l'intervention manuelle de la bibliothécaire. Voyez encore quelques exemples :

Problème : Même titre, auteurs différents, années différentes.

Question : Quelle-est l'année correcte?

Voyez le tableau 16.

Tableau 16 : Quelle-est l'année correcte-1?

AUT: EMBRAPA. Centro Nacional de Pesquisa de Suinos e Aves.
 TIT: Frango de corte Embrapa 021.
 PUB: Concordia: EMBRAPA-CNPSA, [199-].
 ANO: 1990
 UDS: CNPSA;

AUT: EMBRAPA. Centro Nacional de Pesquisa de Suinos e Aves.
 TIT: Frango de corte Embrapa 021.

PUB: Concordia, 1997.

ANO: 1997

UDS: CPAMN-UEPP;

AUT: FIGUEIREDO, E. A. P. de; SCHMIDT, G. S.; ROSA, P. S.; JAENISCH, F. R. F.; SCHEUERMANN, G. N.

TIT: Frango de corte: EMBRAPA 021.

PUB: Concordia: Embrapa-CNPSA, [199-]-.

ANO: 1990

UDS: CNPSA;

Problème : Même titre, auteurs différents, années différentes.

Question : Quelle-est l'année correcte?

Voyez le tableau 17.

Tableau 17 : Quelle-est l'année correcte-2?

AUT: EMBRAPA.

TIT: Wordstar user's guide: for wordstar release 1.0.

PUB: Brasilia: DMQ, 1981

ANO: 1981

UDS: CNPMF;

AUT: EMBRAPA. Departamento de metodos Quantitativos.

TIT: Wordstar user's guide: for wordstar release 1.0.

PUB: Brasilia, 1991

ANO: 1991

UDS: CNPA;

AUT: EMBRAPA. Departamento de Metodos Quantitativos.

TIT: Wordstar user's guide: for wordstar release 1.0.

PUB: Brasilia, 1981.

ANO: 1981

UDS: AI-SEDE;CENARGEN;CPATSA;CPPSE;

Problème : Même titre, auteurs différents, années égales.

Question : Quel registre conserver?

Voyez le tableau 18.

Tableau 18 : Quel registre conserver-1?

AUT: EMBRAPA. Secretaria de Cooperacao Internacional.
TIT: A cooperacao internacional na Embrapa: Documento orientador-versao 3.0.
PUB: Brasilia, 1998.
ANO: 1998
UDS: AI-SEDE;CPATSA;CNPGC;CENARGEN;

AUT: EMBRAPA.
TIT: A cooperacao internacional na Embrapa: documento orientador versao 3.0.
PUB: Brasilia, 1998.
ANO: 1998
UDS: CPAC;CNPGL;

Problème : Même titre, auteurs différents, années égales.

Question : Quel registre conserver?

Voyez le tableau 19.

Tableau 19 : Quel registre conserver-2?

AUT: EMBRAPA. Secretaria do Programa Sistemas de Producao de Materias-Primas.
TIT: Programa sistemas de producao de materias-primas-07.
PUB: Brasili: EMBRAPA-SPI, 1996.
ANO: 1996
UDS: CPAMN-UEPP;

AUT: EMBRAPA. Secretaria Executiva do Programa Sistemas de Producao de Materias-Primas.
TIT: Programa sistemas de producao de materias-primas-07.
PUB: Brasilia: EMBRAPA-SPI, 1996.
ANO: 1996
UDS: CPAMN;

AUT: EMBRAPA. Centro Nacional de Pesquisa de Algodao-.

TIT: Programa sistemas de producao de materias-primas -07.

PUB: Brasilia: Embrapa-SPI, 1996.

ANO: 1996

UDS: CPATC;

AUT: EMBRAPA. Centro Nacional de Pesquisa de Algodao-CNPA. (Campina Grande, PB.

TIT: Programas sistemas de producao de materias-primas,07.

PUB: Brasilia: Embrapa-SPI, 1996.

ANO: 1996

UDS: CPATC;

La difficulté des cas présentés ci-dessus est que si l'algorithme d'identification de doublons est adapté (ou réduit) pour "unifier" les duplicata, les risques **d'identification incorrecte** augmentent. Des références différentes seront identifiées par l'algorithme comme des doublons. Il y a des références avec le même auteur, même titre dans des années différentes qui ne sont pas des doublons. Un exemple est montré tableau 20.

Tableau 20 : Références similaires mais pas des doublons

AUT: ARAI, T.; MITSUI, A.

TIT: Cellular biochemical activities and induction of synchronous growth in a unicellular aerobic nitrogen-fixing cyanobacterium *Synechococcus* sp. strain Miami BG 043511.

PUB: In: Abst. VII International Congress of Photosynthesis, Providence, Rhode Island., v., p.-, 1986.

ANO: 1986

UDS: CNPAB;

AUT: ARAI, T.; MITSUI, A.

TIT: Cellular biochemical activities and induction of synchronous growth in a unicellular aerobic nitrogen-fixing cyanobacterium, *Synechococcus* sp. strain Miami BG 043511.

PUB: In: Progress in Photosynthetic Research, v., p. 649-652, 1987.

ANO: 1987

UDS: CNPAB;

Une autre question non résolue qui nécessite de l'intervention du bibliothécaire est celle des auteurs homographes. Ces références exigent une identification positive du premier nom pour définir à qui attribuer l'article. Ce problème concerne directement l'item "production scientifique" que nous traiterons ensuite. Le plus célèbre se rapporte à une chercheuse déjà décédée, Johanna Dobereiner, mariée avec un autre chercheur de l'Embrapa, Jurgen Dobereiner. La base originale avait 640 références DOBEREINER, J.; huit références pour DOBEREINER, Jurgen; 42 références pour DOBEREINER, Johanna. La base traitée a réduit ce nombre pour 589 références DOBEREINER, J.; neuf références DOBEREINER, Jurgen et 41 références DOBEREINER, Johanna.

Il y a d'autres occurrences similaires: 188 références OLIVEIRA, R. P.; OLIVEIRA, R. P. de.; Qui peuvent être des articles des chercheurs Roberto Pedroso de Oliveira ou Ronaldo Pereira de Oliveira; 43 références WENTZ, I.; Qui peuvent être de l'un des trois frères Wentz, Ivo (20 références individuels), Ilmo (1 registre individuel) ou Ingon (3 références individuels); 139 références PEREIRA, J. R.; Qui peuvent appartenir à trois chercheurs : José Rodrigues Pereira, José Ribamar Pereira ou Joaquim Rezende Pereira.

De la même façon, il n'est pas possible de résoudre automatiquement des erreurs d'indexation comme ceux d'entrer des auteurs comme : FILHO, ?. (61 auteurs); NETO, ?. (59 auteurs); SOBRINHO, ?. (24 auteurs); JUNIOR, ?. (42 auteurs) qui nécessitent être corrigés pour UN TEL FILHO, UN TEL NETO, ?. En outre, nécessitent être corrigés à la main 71 références entrées comme EMATER qui en réalité est seulement la première partie de l'abréviation, l'abréviation générale pour **Empresa de Assistência Técnica e Extensão Rural**. L'abréviation correcte serait EMATER-RJ ou EMATER-MG, en attachant l'abréviation de l'état du Brésil auquel l'Emater appartient.

Le champ affiliation

En profitant des caractéristiques uniques des bases Pascal et WOS, le champ affiliation a été analysé. Une appréciation rapide a indiqué une grande non-conformité dans les citations des adresses et noms d'institutions. Dans la majorité des cas, les noms et adresses de travail sont cités par les chercheurs eux-mêmes. Nous avons choisi, comme exemple, les 90 manières dans lesquelles nom et adresse de l'Embrapa Ressources Génétiques et Biotechnologie (Cenargen) ont été citées dans notre échantillon d'analyse. À observer que toutes les références ont été considérées "uniques" par le programme de traitement mais ont été

identifiées comme étant de la même institution pour avoir dans quelque part le nom "Embrapa" et ainsi ont été réunis :

1. Area de Biologia Celular - Cenargen/Embrapa, C.P. 02372, CEP 70.849-970, Brasilia- DF, Brazil
2. Area de Biologia Celular, Embrapa-Cenargen, C.P. 02372, CEP 70770-900, Brasilia- DF, Brazil
3. Area de Biotecnologia, CENARGEN/EMBRAPA, SAIN, Parque Rural, Caixa Postal 02372, CEP 70770-900, Brasilia-DF, Brazil
4. Area de Biotecnologia, CENARGEN/EMBRAPA, SAIN, Parque Rural, Caixa Postal 02372, CEP 70770-900, Brasilia, DF, Brazil
5. CENARGEN, Parque Estacao Biologica, Av. W/5 Norte, 70770-900, Brasilia, DF, Brazil
6. CENARGEN/ Embrapa, SAIN 70, 770-900, Brasilia, DF, Brazil
7. CENARGEN/ EMBRAPA, SAIN-70, 770-900 Brasilia, DF, Brazil
8. Cenargen/ Embrapa. Sain Parque Rural, 70770-900, Brasilia, D.F., Brazil
9. Cenargen/Embrapa, 70849-970 Brasilia, DF, Brazil
10. CENARGEN/EMBRAPA, CP 02372, 70879-970, Brasilia, DF, Brazil
11. CENARGEN/EMBRAPA, SAIN-Parque Rural-Final W/5 Norte, Brasilia, DF 70770-900, Brazil
12. Cenargen-EMBRAPA, SAIN-Parque Rural CP, 02372 CEP, 70770-970 Brasilia DF, Brazil
13. Centro Nacional de Recursos Geneticos e Biotecnologia, EMBRAPA, 70770-900 Brasilia, DF, Brazil
14. Centro Nacional de Recursos Geneticos e Biotecnologia, Embrapa/Cenargen, Brasilia, Brazil
15. Centro Nacional de Recursos Geneticos e Biotecnologia/EMBRAPA, Brasilia, DF, Brazil
16. Centro Nacional de Recursos Geneticos e Biotecnologia-EMBRAPA, 70770-900 Brasilia, DF, Brazil
17. Centro Nacional de Recursos Geneticos, Embrapa, P.O. Box 02372, Brasilia, DF, Brazil
18. Controle Biologico, EMBRAPA Recursos Geneticos e Biotecnologia (CENARGEN). P.O. Box 02372, Brasilia, D.F., CEP 70770-900, Brasilia, Brazil
19. Embrapa - Genetic Resources and Biotechnology, C.P.02372, 70849-970 Brasilia/DF, Brazil
20. EMBRAPA - Genetic Resources Conservation Area (CENARGEN), CP 02372, CEP 70849-970, Brasilia, DF, Brazil
21. Embrapa - Recursos Geneticos e Biotecnologia-CENARGEN, Parque Estacao Biologica, PqEB - Final Av. W5 Norte, 70770-900 Brasilia-DF, Brazil
22. Embrapa Genetic Resources & Biotechnology (CENARGEN), SAIN-Parque Rural, Caixa Postal 02372, 70770-900 Brasilia, Brazil
23. Embrapa Genetic Resources and Biotechnology (Cenargen), PqEB Final Av. W5 Norte., CEP: 70770-900 Brasilia, DF, Brazil
24. Embrapa Genetic Resources and Biotechnology, Brasilia DF, Brazil
25. EMBRAPA Genetic Resources and Biotechnology, Brasilia, Brazil
26. Embrapa Genetic Resources and Biotechnology, CP 02372, 70849-970 Brasilia DF, Brazil
27. EMBRAPA Recursos Ceneticos e Biotecnologia, P.O. Box 02372, 70770-900, Brasilia, DF, Brazil
28. Embrapa Recursos Genaticos e Biotecnologia, S.A.I.N. Parque Rural, Final W5, Asa Norte, 70770-900, Brasilia-DF, Brazil
29. EMBRAPA Recursos Geneticos e Biotecnologia (CENARGEN), C.P. 02372, Brasilia, D.F., CEP 70770-900, Brazil

30. Embrapa Recursos Geneticos e Biotecnologia (CENARGEN), SAIN, Pq. Rural., C.P. 02372, CEP 70849-970, Brasilia, Brazil
31. Embrapa Recursos Geneticos e Biotecnologia (CENARGEN), SAIN, Pq. Rural., C.P. 02372, CEP 70849-970, Brasilia, DF, Brazil
32. Embrapa Recursos Geneticos e Biotecnologia, Brasilia, DF, 70770-900, Brazil
33. EMBRAPA Recursos Geneticos e Biotecnologia, Brasilia-DF 70770 900, Brazil
34. Embrapa Recursos Geneticos e Biotecnologia, Brazil
35. Embrapa Recursos Geneticos e Biotecnologia, C.P. 02372, 70849-970, Brasilia, DF, Brazil
36. Embrapa Recursos Geneticos e Biotecnologia, C.P. 02372, CEP 70849-970, SAIN, Parque Rural, Brasilia, DF, Brazil
37. EMBRAPA Recursos Geneticos e Biotecnologia, Cx Postal 02372, 70849-970 Brasilia, DF, Brazil
38. Embrapa Recursos Geneticos e Biotecnologia, Cx. Postal 02372, Brasilia - DF, 70849-970, Brazil
39. Embrapa Recursos Geneticos e Biotecnologia, Cx. Postal 02372, Brasilia, DF, 70849- 970, Brazil
40. Embrapa Recursos Geneticos e Biotecnologia, P.O. Box 02372, 70849-970 Brasilia, DF, Brazil
41. EMBRAPA Recursos Geneticos e Biotecnologia, Parque Estacao Biologica, Brasilia, Brazil
42. Embrapa Recursos Geneticos e Biotecnologia, Parque Estacao Biologica, Final Av. W3 Norte, 70770-900 - Brasilia-DF, Brazil
43. EMBRAPA Recursos Geneticos e Biotecnologia, Parque Estacao Ecologica Norte, CP 02372, Final W5 Norte, Brasilia, DF CEP 70.770-900, Brazil
44. Embrapa Recursos Geneticos e Biotecnologia, SAIN Parque Rural, 70770-900 Brasilia (DF), Brazil
45. Embrapa, Cenargen, Caixa Postal 02372, Brasilia, DF, 70770-900, Brazil
46. EMBRAPA, Genetic Resources and Biotechnology - Caixa Postal 02372, Brasilia-DF, 70770-900, BrazilCT- International Scientific Meeting of CBN, 4CL- Salvador, Bahia BRACY- 1998-11-03
47. EMBRAPA/ CENARGEM, CEP 70-770, DF, Brazil
48. EMBRAPA/ CENARGEN, C. Postal 2372, CEP 70770-900, Brasilia DF, Brazil
49. EMBRAPA/CENARGEN C. Postal 02.372, 70 849 970 Brasilia, DF, Brazil
50. Embrapa/Cenargen, C.P 02372, 70849-970 Brasilia-DF, Brazil
51. EMBRAPA/CENARGEN, C.P. 02372, 70849-970 Brasilia, DF, Brazil
52. EMBRAPA/CENARGEN, C.P. 02372, 70849-970, Brasilia, DF, Brazil
53. EMBRAPA/CENARGEN, C.P. 02372, 70849-970, Brasilia., D.F., Brazil
54. EMBRAPA/CENARGEN, Centro Nacional de Pesquisa de Recursos Geneticos e Biotecnologia, SAIN Parque Rural, Brasilia, Brazil
55. Embrapa/Cenargen, CP 02372, 70849-970 Brasilia-DF, Brazil
56. EMBRAPA/CENARGEN, CPO 02372-CEP 70770-900 SAIN Parque Rural, WS Norte Final, Brasilia DF, Brazil
57. EMBRAPA/CENARGEN, P.O. Box 02372, 70.849-970, Brasilia-DF, Brazil
58. EMBRAPA/CENARGEN, SAIN Parque Rural, CP 02372, Brasilia, DF, Brazil
59. EMBRAPA/CENARGEN, SAIN, Parque Rural, 70849-970, Brasilia-DF, Brazil
60. EMBRAPA/CENARGEN/ACB, Brasilia, DF, Brazil
61. EMBRAPA/National Center of Genetic Resources and Biotechnology, SAIN Parque Rural, Brasilia, DF 70770-900, Brazil
62. Embrapa/Recursos Geneticos e Biotecnologia (Cenargen), C. Postal 02372, 70870-900 Brasilia, DF, Brazil

63. Embrapa/Recursos Geneticos e Biotecnologia, Cx. Postal 02372, Brasilia - DF, 70849-970, Brazil
64. EMBRAPA-CENARGEN, P.O. Box 02372, Brasilia, DF, Brazil
65. EMBRAPA-CENARGEN, Parque Rural, Caixa Postal 02372, Brasilia DF, Brazil
66. Embrapa-Cenargen, PqEB- Final Av. W3 Norte, Brasilia, DF70770-900, Brazil
67. Embrapa-Cenargen, Setor de Areas Isoladas Norte Parque Rural, 70, 710-900, Brasilia, Distrito Federal, Brazil
68. Embrapa-Recursos Geneticos e Biotecnologia -Biotechnology Building -Parque Estacao Biologica PqEB, Final W3 Norte, 70770-900 Brasilia-DF, Brazil
69. Embrapa-Recursos Geneticos e Biotecnologia P.O. Box 02372, 70849-970 Brasilia-DF, Brazil
70. Embrapa-Recursos Geneticos e Biotecnologia, Caixa Postal 02372, 70770-900, Brasilia, DF, Brazil
71. EMBRAPA-Recursos Geneticos e Biotecnologia, SAIN Parque Rural, Av. W5 Norte, 70770-900 Brasilia DF, Brazil
72. Empresa Brasileira de Pesquisa Agropecuaria, EMBRAPA- CENARGEN, Sain, Brazil
73. Laboratorio de Espectrometria de Massa EMBRAPA-CENARGEN, P.O. BOX 02372, Brasilia, DF, Brazil
74. Laboratorio de Espectrometria de Massa, Embrapa/Cenargen, Brasilia/DF, 70770-900, Brazil
75. Laboratorio de Espectrometria de Massa, EMBRAPA-CENARGEN, P.O. BOX 02372, Brasilia, DF, Brazil
76. Laboratorio de Genetica de Plantas CENARGEN - EMBRAPA C.P. 02372, Brasilia, DF 70849-970, Brazil
77. Laboratorio de Genetica de Plantas CENARGEN-EMBRAPA - C.P., 02372 Brasilia D.F., Brazil
78. Laboratorio de Genetica de Plantas, CENARGEN-EMBRAPA C.P. 02372, Brasilia D.F. 70849-970, Brazil
79. Laboratorio de Genetica de Plantas, CENARGEN-EMBRAPA C.P. 02372, Brasilia D.F. 79849-970, Brazil
80. Laboratorio de Genetica de Plantas, Embrapa Recursos Geneticos e Biotecnologia C.P.02372, 70849-970 Brasilia DF, Brazil
81. Laboratorio de Genetica de Plantas, Embrapa Recursos Geneticos e Biotecnologia, C.P. 02372, 70849-970, Brasilia, D.F., Brazil
82. Laboratorio de Introducao e Expressao de Genes, Embrapa Recursos Geneticos e Biotecnologia, Parque Estacao Biologica, Final Av. W3 Norte, Brasilia. DF, 70770- 900, Brazil
83. Laboratorio de Microbiologia, EMBRAPA, Brasilia, Brazil
84. Laboratorio de Transferencia de Genes, Embrapa Recursos Geneticos e Biotecnologia, Parque Estacao Biologica (PqEB), 70.849-970, Brazil
85. Laboratory of Molecular Biology, Cenargen-Embrapa, Brasilia, DF, Brazil
86. Microbial Molecular Genetics Laboratory (LGM), EMBRAPA Recursos Geneticos e Biotecnologia (CENARGEN), C.P 02372, Brasilia, D.F., CEP 70770-900, Brazil
87. Microbial Molecular Genetics Laboratory (LGM), EMBRAPA Recursos Geneticos e Biotecnologia (CENARGEN), C.P. 02372, Brasilia, D.F., CEP 70770-900, Brazil
88. National Centre of Genetic Resources and Biotechnology, Cenargen/Embrapa, S.A.I.N. Parque Rural, Final W5, Asa Norte, 70770-900, Brasilia, Brazil
89. National Research Center for Genetic Resources and Biotechnology (CENARGEN/ EMBRAPA), SAIN Parque Rural, Brasilia-DF, Brazil
90. SAIN Parque Rural, Av. W5 Norte, CEP: 70770-900 Brasilia (DF), Brazil

Celle-ci n'est pas une exclusivité de ce centre de recherche. C'est une faute commune pour la plupart des institutions. C'est pour cela qu'avant d'analyser une base, il est nécessaire d'examiner comment les champs ont été remplis et adopter des stratégies de normalisation.

Malheureusement, comme l'a indiqué Quoniam (1996) chaque base est différente et des extraits d'une même base peuvent aussi diverger selon l'année qui est retenue. Ainsi, pour faire des analyses plus précises, les analystes doivent considérer la manque de normalisation dans le champ affiliation. Dans le cas de l'Embrapa cet exemple sert aussi d'alerte. Plus grand étant le nombre de citations normalisées, plus facile il serait pour récupérer des références et aussi de les compter. Un grand nombre d'agences de financement de R&D comme la Fapesp et le CNPq au Brésil et d'autres dans la scène internationale, suivent périodiquement l'évolution de la production scientifique dans ces bases et produisent des indicateurs de ces références. Une filiation normalisée dans les articles soumis à la publication augmente la chance de récupération de tous les articles et peut se transformer aussi en plus de ressources financières ou de prestige pour une institution.

Les priorités des centres de recherche

La base Production scientifique représente, en moyenne, 13,8% de toutes les références "valides" de chaque centre de recherche. Néanmoins, 21 centres de recherche ont leur production scientifique au-dessous de cette moyenne. Les comptages de production scientifique par références "valides" présentent une grande variation. Ils vont de zéro à 34,5%. Les trois dernières ont des scores de 1 à 1,7% et les trois premières de 33,7 à 34,5%, avec un écart-type de 9,8. L'explication la plus plausible pour une telle variance est que dans les centres avec des bas scores, la base Production scientifique, qui fait partie de la Collection documentaire, n'est pas le dépôt central du travail des chercheurs. Ceci est si vrai qu'un centre comme l'Embrapa Manioc et Arboriculture fruitière Tropicale, identifié par l'abréviation CNPMF, qui sustente une significative production scientifique et existait même avant la fondation de l'Embrapa, en 1973, en étant incorporé par elle, compte seulement 195 références dans son catalogue de production scientifique. Cette possibilité très réelle réduit la représentativité de la BDPA comme la collection documentaire de la recherche agronomique de l'Embrapa et pointe des pistes de travail futur. Un tableau de synthèse est fourni tableau 21. Il est ordonné par le pourcentage de la production scientifique sur les références valides. Les centres signalés en gris sont au-dessous de la moyenne.

Tableau 21 : Références valides par pourcentage de la Production scientifique, avril 2005

Centre	total reg.	total valides	prod. scient.	prod. scient. valides	%
CNPSO	23375	22824	7939	7884	34,5
CNPDIA	3997	3799	1357	1308	34,4
CNPMS	16413	15952	5658	5379	33,7
CNPM	1173	1164	382	373	32
CTAA	6147	4439	1375	1340	30,2
AI-SEDE	29017	27377	8189	8070	29,5
CPAC	24192	22883	5223	5094	22,2
CNPGC	9112	7713	1951	1716	22,2
CNPMA	4261	3637	793	759	20,8
CNPH	28722	23943	4939	4731	19,7
CPAA	10377	9614	1935	1732	18
CNPAF	21395	13131	2373	2349	17,9
CPPSE	14627	14622	2565	2565	17,5
CNPC	16664	7989	1367	1361	17
CPATU	32749	31543	4855	4750	15
CNPTIA	8552	7031	1055	1000	14,2
CNPS	10998	7872	1110	1099	13,9
CPATSA	28326	28183	3905	3881	13,8
CNPSA	14098	6147	2250	715	11,6
CPAMN	17692	17497	1946	1928	11
CNPF	30588	28428	3087	3055	10,7
CENARGEN	22503	19318	1938	1902	9,8
CPAF-AC	8498	8242	840	805	9,7
CPAF-RO	8664	8343	790	784	9,4
CNPA	16120	13526	1283	1243	9,2
CPACT	7704	7644	684	684	8,9
CPAF-RR	5303	3086	267	267	8,6
CPAP	26908	26082	1530	1503	5,8
CPAO	21964	21004	1047	1041	4,9
CPAF-AP	7325	7234	357	355	4,9
CNPAT	8052	7631	321	311	4,1
CNPUV	5683	5667	213	213	3,7
CPATC	18066	17678	527	525	2,9
CNPAB	30588	28691	1918	770	2,7
CNPGL	13543	10105	181	177	1,7
CPAMN-UEPP	10693	10626	152	151	1,4
CNPMF	19835	18956	269	195	1
CNPT	0	0	0	0	0
CPPSUL	0	0	0	0	0
Total	583924	519621	76571	72015	13,8

Note. total reg. = nombre total de références dans la Collection documentaire de l'Embrapa.
total valides = nombre de références sélectionnés pour apparaître dans la Collection documentaire de l'Embrapa.
prod. scient. = nombre de références de la Production scientifique du centre de recherche.
valides = nombre de références sélectionnées pour apparaître dans la Production scientifique du centre de recherche.

Les limites des bases collection documentaire et production scientifique

Les limites de la solution automatique bibliométrique sont aussi les propres limites de la BDPA. D'abord, il y a la question déjà abordée de que la base Production scientifique n'a pas la même priorité d'indexation en tous les centres de recherche (tableau 25).

D'autres questions importantes sont, d'un côté, la manque de listes actuelles d'institutions, de journaux et revues scientifiques, de dates (années) et titres. Une liste de titres, par exemple, si elle était disponible pendant le catalogage, éviterait plusieurs erreurs d'indexation et d'orthographe pendant l'indexation de documents par d'autres centres de recherche, qui peut résulter en doublons comme les exemples cités auparavant. L'existence de listes précises d'auteurs et titres est d'ailleurs un des critères de qualité de bases de données du SCOUG (Rittberger et Rittberger, 1997 ; Basch, 1995 ; Medawar, 1995 ; Quoniam, 1996, Faria, 2001).

Un autre critère de qualité (Jacsó, 1993b ; Mintz, 1990 ; O'Neill et Vizine-Goetz, 1988 ; Sule, 1986) absent est un mécanisme central d'indexation pour filtrer des éventuelles duplications, en critiquant les nouvelles références avant qu'ils soient introduits dans la base.

Il manque aussi un manuel de procédures pour le catalogage de références pour guider les nouveaux bibliothécaires. Ce document est important pour éclaircir des éventuels doutes qui en aval deviennent significatifs puisqu'ils génèrent des doublons. C'est d'ailleurs un autre critère de qualité SCOUG (Medawar, 1995 ; Basch 1995, Faria, 2001). Toutes ces questions amèneraient à une plus grande convergence des pratiques d'indexation.

Ces limitations s'accroissent du fait que l'indexation des références doit être faite d'une manière décentralisée, dans les 39 bibliothèques séparées dans l'espace et autonomes dans l'administration de chaque sous-base.

Pour augmenter la fiabilité des bases Embrapa et permettre son utilisation comme un instrument d'accompagnement des plans annuels de travail des centres de recherche de l'Embrapa, la Secrétaire d'Administration et Stratégie (SGE), en collaboration avec la CP-Ainfo a adopté des mesures pour encourager l'indexation de la production des centres de recherche et la garde permanente de ces documents dans les bibliothèques. Une autre mesure

importante a été la création de champs spécifiques pour indexer ces documents dans la BDPA. Ces champs permettront dans le futur, la génération automatique d'indicateurs et aussi l'accompagnement de gestion par les systèmes d'administration stratégique de l'Embrapa. Ces mesures sont prévues d'être opérationnelles au cours des deux prochaines années. Les nouveaux champs créés sont les suivants :

"Type de Production Technique et Scientifique" avec les options suivantes : Article en Périodique Indexé ; Chapitre en Livre Technique et Scientifique ; Article dans des Actes de Congrès / Note Technique ; Résumé dans des Actes de Congrès et Orientation de Thèse de Doctorat ;

"Type de Production de Publications Techniques" avec les options suivantes : Circulaire Technique ; Communiqué Technique ; Bulletin de Recherche et Développement ; Documents ; Organisation/Édition de Livres ou de Système de Production ; Article de Divulgence dans les Mass Médias.

Les limites du logiciel Ainfo

Nous avons trouvé plusieurs limitations du logiciel Ainfo créé pour l'administration des bases Embrapa. Des listes d'institutions et l'algorithme de critique des indexations sont très importantes parce que les 39 bibliothèques de l'Embrapa, comme a été dit, sont séparées dans l'espace et autonomes dans l'administration et l'indexation de nouvelles références de leur base.

À partir du travail de Corte et al. (2002) nous avons rencontré aussi plusieurs caractéristiques de logiciels de bibliothèques "souhaitables" pour l'Ainfo. Elles sont divisées en six rubriques :

Traitement technique

- Construction de listes (rémissibles) pour auteurs et sujets ;
- Correction automatique de tous les références associées à un auteur ou sujet par l'actualisation des listes d'autorités ou du thésaurus ;
- Cadastre des champs de la connaissance selon la liste de champs de la connaissance du CNPq.

Circulation et référence

- Émission de rapports d'emprunt, sujets les plus consultés dans une période, usager avec plus d'emprunts ;
- Possibilité d'identifier tous les emprunts faits d'une oeuvre déterminée, par période, en ordre alphabétique d'usager.

Récupération d'informations

- Recherche automatique par le thésaurus ;
- Recherche interactive à partir de la sélection du thésaurus ;
- Capacité d'ordonner et classifier les documents recherchés par titre ;
- Capacité d'ordonner et classifier les documents recherchés par sujet ;
- Capacité d'ordonner et classifier les documents recherchés par importance ;
- Capacité d'ordonner et classifier les documents recherchés par type de document ;
- Capacité d'ordonner et classifier les documents recherchés par date ;
- Émission de statistiques de consultations à la base avec nombre d'accès par tranches d'heure, jour, jour de la semaine, jour du mois, mois de l'année, entre autres.

Dissémination de l'information

- Cadastre du profil de l'usager par sujets d'intérêt ;
- Dissémination sélective de l'information.

Processus d'administration

- Création et impression de rapports en références bibliographiques par disciplines, avec nombre d'appel et exemplaires ;
- Création et impression de rapports en références bibliographiques par nombre de titres et exemplaires par champs de la connaissance du CNPq.

Méthode de conversion rétrospective adoptée

- Clés uniques (ISSN, ISBN) ;
- Digitalisation (scanner) ;
- En ligne (recherche dans des catalogues coopératifs).

L'analyse de la production scientifique de l'Embrapa

Prenant comme base les matrices générées par Penteado et al. (2002) - année de publication/auteur, année de publication/affiliation, auteur/affiliation - et en fonction des limitations et du haut nombre d'auteurs avec fréquence unique ou double, nous avons décidé de comparer les résultats préliminaires des analyses bibliométriques en croisant les champs extraits des bases Collection documentaire et Production scientifique de l'Embrapa pendant une période spécifique de temps, les années 1998, 1999, 2000, 2001 et 2002. Nous avons décidé aussi, pour faire des comparaisons, de réaliser des analyses bibliométriques à partir de corpus constitué à partir de bases internationales, comme CAB, Agricole, Pascal et World of Science (WOS), dans cette même période.

Méthodologie

La base Collection documentaire nettoyée et reformatée a été importée dans la banque de données Microsoft Access. Nous avons appliqué un filtre pour les années de l'étude déjà citées. Les 54.343 références trouvées ont été réexportées pour le format .CSV, reformatés dans l'Infotrans pour le format bibliométrique et importés par le logiciel VantagePoint qui a généré les matrices année de publication/auteur, année de publication/affiliation et auteur/affiliation.

Une enquête a été réalisée dans la base Production scientifique de l'Embrapa originelle pour la période étudiée. Les résultats (23.678 références) ont été soumis au programme de nettoyage et reformatage d'Infotrans. Ensuite, la base nettoyée et reformatée a été importée par VantagePoint. Une vérification préliminaire a indiqué la présence des années de publication différentes de celles étudiées. Un filtre a alors été appliqué en sélectionnant seulement des années 1998, 1999, 2000, 2001 et 2002. Avec les 19.371 références résultantes ont été générés les matrices année de publication/auteur, année de publication/affiliation et auteur/affiliation.

Quand nous avons travaillé avec des bases internationales, l'expression initiale de recherche a été : Dans les années 1998, 1999, 2000, 2001 et 2002, toutes les références avec EMBRAPA en affiliation.

Une deuxième recherche a été faite avec l'expression suivante : Dans les années 1998, 1999, 2000, 2001 et 2002, toutes les références avec EMBRAPA en affiliation ou CPAFAC ou

CNPAB ou CTAA ou CNPAT ou CPAO ou CNPA ou CPAFAP ou CPAA ou CPATU ou CNPAF ou SAPC ou CNPC ou CPAC ou CPACT ou SCT ou CNPF ou CNPGC ou CNPGL ou CNPH ou CNPTIA ou CNPDIA ou CNPMF ou CNPMA ou CPAMN ou CNPMS ou CNPM ou SNT ou CPAP ou CPPSE ou CPPSUL ou CENARGEN ou CPAFRO ou CPAFRR ou CPATSA ou CNPS_o ou CNPS ou CNPSA ou CPATC ou CNPT ou CNPUV. Cette deuxième expression de recherche a retourné beaucoup de références d'organisations différentes de l'Embrapa et au contraire de l'attendu, n'a pas augmenté le nombre de nouvelles références. En conséquence, nous avons suivi Quoniam (1996) et adopté le principe de la parcimonie pour maintenir l'expression de recherche la plus simple.

Nous avons noté ensuite que les bases CAB et Agricole même en ayant une ample couverture de l'Embrapa, exigeaient une très grande charge de travail manuel pour réaliser des études bibliométriques. Elles ne donnaient pas d'options pour l'exportation des résultats de recherche. Il serait nécessaire de réaliser les recherches, copier et coller le résultat de chaque recherche dans un fichier séparé, des fonctions réalisées par des robots de recherche sur l'Internet. Comme l'Embrapa ne disposant pas encore de ces instruments automatiques de récupération d'informations sur l'Internet, nous avons décidé de ne pas travailler avec ces deux bases.

Nous avons aussi profité de la forme de construction des bases WOS et Pascal, avec des affiliations de tous les auteurs, pour étudier des sujets comme la coopération internationale et nationale entre les organismes scientifiques partenaires de l'Embrapa. Ont été récupérées 868 références dans la base Pascal et 2.421 références dans la base WOS. Ils ont été reformatés et nettoyés sous Infotrans. Un registre de la base WOS est montré dans le tableau 22.

Tableau 22 : Le registre WOS original

PT J
AU Angelocci, LR
Nova, NAV
Coelho, MA
Marin, FR
TI Measurements of net radiation absorbed by isolated acid lime trees (<i>Citrus latifolia</i> Tanaka)
SO JOURNAL OF HORTICULTURAL SCIENCE & BIOTECHNOLOGY
LA English
ID APPLE TREE; TRANSPIRATION; INTERCEPTION
C1 Univ Sao Paulo, Escola Super Agr Luiz Queiroz, Dept Ciencias Exatas, BR-13418900 Piracicaba, SP, Brazil.
EMBRAPA, Ctr Nacl Pesquisa Mandioca & Fruticultura, Cruz Das Alamas, BA, Brazil.
EMBRAPA, Ctr Nacl Pesquisa Monitoramento Satelite, Campinas, SP, Brazil.
RP Angelocci, LR, Univ Sao Paulo, Escola Super Agr Luiz Queiroz, Dept Ciencias Exatas, CP 9, BR-13418900 Piracicaba, SP, Brazil.
EM lrangelo@carpa.ciagri.usp.br
PU HEADLEY BROTHERS LTD
PI ASHFORD
PA INVICTA PRESS, ASHFORD TN24 8HH, KENT, ENGLAND
PD SEP
PY 2004
VL 79
IS 5
BP 699
EP 703
SC Horticulture
UT ISI:000224361600006
ER

Le tableau 23 montre le même registre de la base WOS après le traitement.

Tableau 23 : Le registre WOS traité

AU: ANGELOCCI LR; NOVA NAV; COELHO MA; MARIN FR
 TI: Measurements of net radiation absorbed by isolated acid lime trees (*Citrus latifolia* Tanaka)
 FI: Univ Sao Paulo, Escola Super Agr Luiz Queiroz, Dept Ciencias Exatas, BR-13418900 Piracicaba, SP, Brazil.; EMBRAPA, Ctr Nacl Pesquisa Mandioca & Fruticultura, Cruz Das Alamas, BA, Brazil.; EMBRAPA, Ctr Nacl Pesquisa Monitoramento Satelite, Campinas, SP, Brazil.
 JN: JOURNAL OF HORTICULTURAL SCIENCE & BIOTECHNOLOGY
 AP: 2004
 LA: English
 ED: APPLE TREE; TRANSPIRATION; INTERCEPTION

Le tableau 24 montre un registre de la base Pascal.

Tableau 24 : Le registre Pascal original

DA- 15824153|
 AN- 02-0541326|
 ET- Lignin characteristics and density fractions of termite nests in an Amazonian rain forest: indicators of termite feeding guilds?|
 AU- AMELUNG Wulf^MARTIUS Christopher^BANDEIRA Ademar G^GARCIA Marcos V B^ZECH Wolfgang|
 CS- Institute of Soil Science and Soil Geography, University of Bayreuth, 95440 Bayreuth, Germany^Center for Development Research (ZEF), Walter-Flex-Str. 3, 3113 Bonn, Germany^Departamento de Sistemática e Ecologia, Centro de Ciências Exatas e da Natureza, Universidade Federal da Paraíba, 58059-000 João Pessoa, PB, Brazil^Brazili Embrapa Amazonia Ocidental, c.p. 319, 69.011-970 Manaus, AM, Brazil|
 JN- Soil biology & biochemistry|
 PD- 2002|
 VO- 34|
 NO- 3|
 PG- 367-372|
 SN- 0038-0717|
 CD- SBIOAH|
 AV- INIST-14360|
 RF- 31 ref.|
 DT- P- Serial^A- Analytic|
 CP- United Kingdom|

LA- English|

EA- High termite abundance and diversity in rainforests results in the creation of different termite nests (..) guilds of the studied taxa.|

SC- 002A14B04B|

ED- Lignin^Density distribution^Termite mound^Nest^Foraging behavior^Xylophagous^Epiphyte^Arboricole^Guild^Nasutitermes^Tropical rain forest^Phenols^Organic matter^Brazil^Amazon Basin|

FD- Lignine^Distribution densite^Termitiere^Nid^Comportement approvisionnement^Xylophage^Epiphyte^Arboricole^Guilde^Nasutitermes^Foret ombrophile tropicale^Phenols^Matiere organique^Bresil^Bassin Amazone^Ferralsol xanthique|

SD- Lignina^Distribucion densidad^Termitero^Nido^Conducta abastecimiento^Xilofago^Epifito^Arboricola^Nasutitermes^Bosque tropical humedo^Fenoles^Materia organica^Brasil^Cuenca Amazonas|

BT- Termitidae; Isoptera; Insecta; Arthropoda; Invertebrata; South America; America; Soils; Termite; Feeding behavior; Tropical zone; Ferralsols^Termitidae; Isoptera; Insecta; Arthropoda; Invertebrata; Amerique du Sud; Amerique; Sol; Termite; Comportement alimentaire; Zone tropicale; Ferralsol^Termitidae; Isoptera; Insecta; Arthropoda; Invertebrata; America del sur; America; Suelo; Termita; Conducta alimenticia; Zona tropical; Ferralsols||

Le tableau 25 montre le même registre de la base Pascal après le traitement avec Infotrans.

Tableau 25 : Le registre Pascal traité

AU: AMELUNG WULF; MARTIUS CHRISTOPHER; BANDEIRA ADELMAR G; GARCIA MARCOS VB; ZECH W

TI: Lignin characteristics and density fractions of termite nests in an Amazonian rain forest: indicators of termite feeding guilds?

FI: Institute of Soil Science and Soil Geography, University of Bayreuth, 95440 Bayreuth, Germany; Center for Development Research (ZEF), Walter-Flex-Str. 3, 53113 Bonn, Germany; Departamento de Sistemática e Ecologia, Centro de Ciências Exatas e da Natureza, Universidade Federal da Paraíba, 58059-000 João Pessoa, PB, Brazil; Brazi Embrapa Amazonia Ocidental, c.p. 319, 69.011-970 Manaus, AM, Brazil

JN: Soil biology & biochemistry (United Kingdom)

AP: 2002

LA: English

PC: Lignin; Density distribution; Termite mound; Nest; Foraging behavior; Xylophagous; Epiphyte; Arboricole; Guild; Nasutitermes; Tropical rain forest; Phenols; Organic matter; Brazil; Amazon Basin

Le travail de nettoyage et reformatage d'auteurs qui a été fait pour les bases Embrapa n'a pas été fait dans la même extension que dans les bases Embrapa. Nous avons listé les auteurs par

ordre alphabétique et les avons consolidé par le critère de proximité géographique. Ce critère n'a pas permis de rencontrer des doublons de chercheurs comme, par exemple : REIS, V. M. et MASSENA REIS, V. Mais nous avons cherché appui en Quoniam (1996, pp.71-72). Il argumente qu' "il est plus judicieux de laisser des erreurs flagrantes et montrer ainsi le côté relatif de tout comptage. Ces erreurs n'introduisent pas globalement d'écart d'interprétation".

Résultats

En considérant les limitations de la base, des politiques d'indexation des centres de recherche et du programme de nettoyage et reformatage, le tableau 26 montre l'évolution de la production scientifique de 36 centres de recherche de l'Embrapa et de son Siège dans la période étudiée. Les bases Embrapa fonctionnent en parallèle avec les bibliothèques. Comme l'Embrapa Café (CAFE) n'a pas de bibliothèque, il n'apparaît pas. Deux autres centres, l'Embrapa Trigo (CNPT) et l'Embrapa Élevage du Bétail Sud (CPPSUL) n'utilisent pas l'Ainfo et pourtant ne font pas partie des bases Embrapa. Les trois apparaissent dans les statistiques des bases internationales. A remarquer encore qu'un centre de recherche, l'Embrapa Moyen-Nord (CPAMN) apparaît deux fois (CPAMN et CPAMN-UEPP) parce qu'il a deux bibliothèques. Dans les bases internationales il apparaît seulement une fois. Les centres de recherche en gris sont ceux-ci au-dessous de la moyenne de 13,8% de la production scientifique par le nombre totale de références dans sa Collection documentaire.

Tableau 26 : Production scientifique de l'Embrapa de 1998 à 2002, par centre

Base Production Scientifique 1998_2002						
Centre	2002	2001	2000	1999	1998	Total
CNPSO	550	435	467	616	374	2442
AI-SEDE	311	257	353	302	326	1549
CPATSA	316	313	230	333	209	1401
CPATU	188	303	371	382	425	1669
CNPH	258	263	233	275	258	1287
CNPMS	461	186	382	157	227	1413
CPAC	30	280	318	366	242	1236
CNPF	192	142	243	75	143	795
CPPSE	191	180	203	131	135	840
CPAMN	118	114	132	161	154	679
CNPGC	53	76	133	158	139	559
CNPAF	166	112	82	133	80	573
CTAA		177	159	143	135	614
CPAP	128	72	161	66	67	494
CENARGEN	149	158	83	86	29	505
CNPTIA	143	126	73	45	56	443
CPAA	75	62	32	135	125	429
CNPA	46	55	76	96	109	382
CNPC	51	42	52	86	59	290
CPAO	29	65	87	85	65	331
CPACT	70	53	56	31	53	263
CNPS	24	30	30	90	79	253
CNPAB	32	28	55	56	64	235
CNPDIA	125	5	24	9	7	170
CNPMA	8		5	42	117	172
CPAF-RO	9	12	13	48	53	135
CPAF-AC	11	15	30	23	43	122
CPAF-AP	12	18	11	50	22	113
CPATC	6	33	10	7	34	90
CNPM	20	21	21	17	6	85
CNPSA	5	4	3	31	35	78
CNPAT	1	2	28	24	23	78
CNPUV	16	8	14	12	13	63
CPAF-RR			3	2	38	43
CNPGL	1	14	5	15	4	39
CPAMN-UEPP			9	9	9	27
CNPMPF		8	14	1		23
Total	3795	3669	4201	4298	3957	19920

Comparaison des bases Collection documentaire et Production scientifique

Quand nous comparons les résultats des bases Production scientifique 98_02 et Collection documentaire (un extrait des années 98_02), des chercheurs comme COSTA, N. de L. et TOWNSEND, C. R. (CPAF-RO); BELTRAO, N. E. de M. (CNPA); ANDRADE JUNIOR, A. S. de. (CPAMN) qui appartiennent à des unités dont les scores se situent au-dessous de la moyenne pour leurs bases Production scientifique sont mieux placés dans la base Collection documentaire. De la même façon, des chercheurs comme JUNQUEIRA, N. T. V.; GASPAROTTO, L.; ALENCAR, J.A de.; MIRANDA, E. E. de.; Respectivement, du CPAC, CPAA, CPATSA et CNPM ont enregistré une croissance de leur production. A remarquer aussi le surgissement dans la base Collection documentaire de chercheurs qui, par définition, seraient au dehors de la base Production scientifique parce qu'ils n'appartiennent pas à l'Embrapa : KAMIKOGA, M. K. (ESALQ); MATSUMOTO, M. N. (UEL); ALBERINI, J. L. (FT-SEMENTES); ABATI, C. (MONSANTO). Les chercheurs en gris dans le tableau 27 appartiennent à des centres au-dessous de la moyenne de 13,8% de la production scientifique par le nombre totale de références dans sa Collection documentaire.

Le tableau 27 indique aussi que deux centres de recherche, Embrapa Soja (CNPSSO) et Embrapa Maïs et Sorgho (CNPMS), dominent pratiquement la scène en termes de quantité d'articles et d'autres travaux scientifiques publiés et indexés dans la base. Ils sont marqués en gras. Les chercheurs des institutions collaboratrices qui sont cités travaillent aussi dans le thème soja. En résumant, la concentration dans des centres spécifiques implique aussi en concentration de thèmes. Ceci se reflète dans le fait que 28 des 55 chercheurs avec la plus grande production scientifique appartiennent à ces deux centres. Comme le champ "auteur" n'était pas obligatoire, il pouvait ne pas être rempli lors de l'étape d'indexation les 1.214 Champs vides qui n'étaient pas dénombrés dans la base originelle, occupent la première place entre les auteurs dans la base Collection documentaire.

Tableau 27 : Auteurs, bases Production scientifique et Collection documentaire 1998/2002

Base Production scientifique traitée 1998_2002			Base Collection documentaire traitée 1998_2002				
# Occurrences	Auteur	Centre	# Occurrences	Auteur	Centre		
1	391	ALMEIDA, L. A. de	CNPSO	1	1214	ChampVide	ChampVide
2	348	KIIHL, R. A. de S	CNPSO	2	424	ALMEIDA, L. A. de	CNPSO
3	298	SANTOS, M. X. dos	CNPMS	3	371	KIIHL, R. A. de S	CNPSO
4	213	FRANCA NETO, J. de B	CNPSO	4	275	SANTOS, M. X. dos	CNPMS
5	205	CARVALHO, H. W. L. de	CPATC	5	208	FRANCA NETO, J. de B	CNPSO
6	200	HUNGRIA, M	CNPSO	6	204	EMBRAPA	SIÈGE
7	198	KRZYZANOWSKI, F. C	CNPSO	7	193	KRZYZANOWSKI, F. C	CNPSO
8	193	CARDOSO, M. J	CPAMN	8	191	CARDOSO, M. J	CPAMN
9	186	HENNING, A. A	CNPSO	9	191	CARVALHO, H. W. L. de	CPATC
10	172	ALENCAR, M. M. de	CPPSE	10	188	HUNGRIA, M	CNPSO
11	169	YORINORI, J. T	CNPSO	11	185	HENNING, A. A	CNPSO
12	165	COSTA, N. P. da	CNPSO	12	184	YORINORI, J. T	CNPSO
13	156	PRIMAVESI, O. M. A. S. P. R. B	CPPSE	13	169	ALENCAR, M. M. de	CPPSE
14	153	OLIVEIRA, A. C. de	CNPMS	14	167	COSTA, N. P. da	CNPSO
15	147	LEAL, M. de L. da S	CPATC	15	156	JUNQUEIRA, N. T. V	CPAC
16	137	SERENO, J. R. B	CPAP	16	154	BELTRAO, N. E. de M	CNPA
17	136	FARIA, L. C	CNPAF	17	153	PRIMAVESI, O. M. A. S. P. R. B	CPPSE
18	132	PRIMAVESI, A. C. P. de A	CPPSE	18	152	LEAL, M. de L. da S	CPATC
19	130	ARANTES, N. E	CNPSO	19	150	SERENO, J. R. B	CPAP
20	129	FREITAS, A. R. de	CPPSE	20	144	OLIVEIRA, A. C. de	CNPMS
21	129	PARENTONI, S. N	CNPMS	21	142	GASPAROTTO, L	CPAA
22	129	CRUZ, I	CNPMS	22	142	RESENDE, G. M. de	CPATSA
23	128	GAZZIERO, D. L. P	CNPSO	23	139	FARIA, L. C	CNPAF
24	124	SOUZA, P. I. M	CPAC	24	138	SOUZA, P. I. M	CPAC
25	124	SILVA, J. F. V	CNPSO	25	137	PRIMAVESI, A. C. P. de A	CPPSE
26	124	BELTRAO, N. E. de M	CNPA	26	135	ARANTES, N. E	CNPSO
27	121	AVILA, A. C. de	CNPH	27	135	HAJI, F. N. P	CPATSA
28	119	RESENDE, G. M. de	CPATSA	28	134	CRUZ, I	CNPMS
29	118	LOPES, C. A	CNPH	29	133	SILVA, J. F. V	CNPSO
30	116	GAMA, E. E. G	CNPMS	30	132	LOPES, C. A	CNPH
31	113	HAJI, F. N. P	CPATSA	31	130	GAZZIERO, D. L. P	CNPSO
32	113	JUNQUEIRA, N. T. V	CPAC	32	128	KAMIKOGA, M. K	ESALQ
33	112	PACHECO, C. A. P	CNPMS	33	127	MATSUMOTO, M. N	UEL
34	109	QUEIROZ, M. A. de	CPATSA	34	126	QUEIROZ, M. A. de	CPATSA
35	107	GIORDANO, L. de B	CNPH	35	123	AVILA, A. C. de	CNPH
36	107	VOLL, E	CNPSO	36	121	ALENCAR, J. A. de	CPATSA
37	106	FARIAS, J. R. B	CNPSO	37	121	COSTA, N. de L	CPAF-RO
38	104	GASPAROTTO, L	CPAA	38	121	NOGUEIRA, A. R. A	CPPSE
39	103	BARBOSA, F. R	CPATSA	39	119	FREITAS, A. R. de	CPPSE
40	103	ASSAD, E. D	CNPTIA	40	119	PARENTONI, S. N	CNPMS
41	102	TORRES, E	CNPSO	41	115	MIRANDA, E. E. de (Coord	CNPM
42	102	NOGUEIRA, A. R. A	CPPSE	42	112	ALBERINI, J. L	FT-SEM
43	101	BORKERT, C. M	CNPSO	43	111	ABATTI, C	MONSAN
44	99	LAMEIRA, O. A	CPATU	44	111	BARBOSA, F. R	CPATSA
45	98	ALENCAR, J. A. de	CPATSA	45	111	RIBEIRO, V. Q	CPAMN
46	97	PEREIRA, J. E	CNPSO	46	109	ASSAD, E. D	CNPTIA
47	96	RIBEIRO, V. Q	CPAMN	47	108	GIORDANO, L. de B	CNPH
48	96	CAMPO, R. J	CNPSO	48	108	MAROUELLI, W. A	CNPH
49	96	PAIVA, E	CNPMS	49	107	TOWNSEND, C. R	CPAF-RO
50	94	MAROUELLI, W. A	CNPH	50	106	VOLL, E	CNPSO
51	94	MAGALHAES, P. C	CNPMS	51	105	FARIAS, J. R. B	CNPSO
52	93	SFREDO, G. J	CNPSO	52	105	GAMA, E. E. G	CNPMS
53	92	ARIAS, C. A. A	CNPSO	53	103	ANDRADE JUNIOR, A. S. de	CPAMN
54	92	NASCIMENTO, W. M	CNPH	54	103	BORKERT, C. M	CNPSO
55	92	BRIGHENTI, A. M	CNPSO	55	101	PEREIRA FILHO, I. A	CNPMS

Comparaison des bases Pascal et WOS

De l'autre côté, une classification différente est apparue dans les bases WOS et Pascal. Les bases internationales ont présenté une plus grande variété. Seulement sept d'entre les 55 premiers auteurs de la base WOS appartenaient à Embrapa Soja (CNPSO) ou à Embrapa Maïs et Sorgho (CNPMS). Les noms de ces centres sont marqués en gras. Dans la base Pascal ce nombre a été encore plus faible, trois des 55 premiers. Pendant que dans les bases de Production scientifique et Collection documentaire 98_02 les 55 premiers chercheurs appartenaient à 14 centres de recherche, ce nombre s'est accru pour 16 centres dans la base Pascal et 17 centres dans la base WOS. Le plus intéressante est que les nouveaux centres qui ont surgi sont dans la majorité entre ceux avec des politiques déficientes d'indexation de leur production scientifique dans les bases Embrapa.

Dans la base WOS, par exemple, si on considère les dix auteurs avec le plus de références, le premier, Dobereiner, J. est de l'Embrapa Agrobiologie (CNPAB), le deuxième, Rigden, D. J., le sixième, Bloch, C. et le septième, De Sa, M. F. G. sont de l'Embrapa Ressources Génétiques et Biotechnologie (Cenargen) et le dixième, Verneque, R. D. est de l'Embrapa Bovins de Lait (CNPGL) trois unités qui sont au-dessous de la moyenne dans les bases Embrapa. Entre les dix, seulement un, Dos Santos, M. X., de l'Embrapa Maïs et Sorgho (CNPMS) est des unités qui "dominaient la scène" dans les bases Production scientifique et Collection documentaire 98_02.

Dans la base Pascal, un nouveau changement. L'Embrapa Ressources Génétiques et Biotechnologie (CENARGEN) sort du podium et tombe à la 26^{ème} place. L'Embrapa Instrumentation Agricole (CNPDIA) place le premier Mattoso, L. H. C. et le dixième, Vaz, C. M. P., l'Université Fédérale du Paraná (UFPR) le sixième et le septième, Mathias, A. L. et Ramos, A. P. et l'Université de British Columbia le cinquième, Saddler, John N. Autres cinq centres divisent avec le CNPDIA les dix premières places : Embrapa Riz et Haricot (CNPAF), Embrapa Amazonie Occidentale (CPAA), Embrapa Agro-industrie d'Aliments (CTAA), Embrapa Élevage du Bétail Sud-Est (CPPSE) et Embrapa Soja (CNPSO), le seul des deux centres "dominateurs de la scène" dans les bases Embrapa à apparaître entre les dix premiers. Autre centre qui apparaît avec force, en spécial, dans la base Pascal est l'Embrapa Agrobiologie (CNPAB). Les centres au-dessous de la moyenne dans la base Production scientifique de l'Embrapa sont signalés en gris dans le tableau 28.

Tableau 28 : Auteurs, bases WOS et Pascal, 1998/2002

Base WOS				Base Pascal			
# Occurrences	Auteur	ClaPa	Centre	# Occurrences	Auteur	ClaWOS	Centre
1	27 DOBEREINER J	17	CNPAB	1	22 MATTOSO LHC	3	CNPDIA
2	27 RIGDEN DJ	205	CENARGEN	2	19 FAGERIA NK	4	CNPAF
3	25 MATTOSO LHC	1	CNPDIA	3	17 ZECH W	14	CPAA
4	25 FAGERIA NK	2	CNPAF	4	15 ASCHERI JLR	sc	CTAA
5	22 NOGUEIRA ARA	9	CPPSE	5	15 SADDLER JOHN N, ED	sc	UBRITCOL
6	22 BLOCH C	45	CENARGEN	6	15 MATHIAS AL	sc	UFPR
7	21 DE SA MFG	sc	CENARGEN	7	15 RAMOS LP	sc	UFPR
8	20 DOS SANTOS MX	970	CNPMS	8	14 HUNGRIA M	13	CNPISO
9	20 VALADARES SD	sc	CPATSA	9	14 NOGUEIRA ARA	5	CPPSE
10	19 VERNEQUE RD	303	CNPGL	10	14 VAZ CMP	23	CNPDIA
11	17 NOBREGA JA	15	UFSCAR	11	13 MARTIN-NETO L	55	CNPDIA
12	17 DA SILVA JFC	sc	CPATSA	12	12 SCHROTH G	26	UHAMB
13	17 HUNGRIA M	8	CNPISO	13	10 BOITEUX LS	25	CNPB
14	17 ZECH W	3	CPAA	14	10 CRESTANA S	84	CNPDIA
15	15 DE ARAUJO GGL	sc	CPATSA	15	10 NOBREGA JA	11	UFSCAR
16	14 CECON PR	sc	UFV	16	9 URQUIAGA S	28	CNPAB
17	14 PAIVA E	22	CNPMS	17	9 DOBEREINER J	1	CNPAB
18	14 PEREIRA JC	766	CNPAB	18	9 AYARZA MA	112	CIAT
19	14 REIS VM	23	CNPAB	19	9 VAN RAIJ B	91	CNPMA
20	14 CRUZ CD	sc	UFV	20	9 MACEDO A	128	CNPDIA
21	14 SIGNORETTI RD	sc	CNPISA	21	8 BALIGAR VC	72	ARS-USDA
22	14 GUIDONI AL	sc	CNPISA	22	8 PAIVA E	17	CNPMS
23	14 VAZ CMP	10	CNPDIA	23	8 REIS VM	19	CNPAB
24	13 PRABHU AS	120	CNPAF	24	8 VILELA L	95	CPAC
25	13 BOITEUX LS	13	CNPB	25	8 WANG SH	sc	UFRRJ
26	13 SCHROTH G	12	UHAMB	26	7 GROSSI DE SA MF	sc	CENARGEN
27	13 GRATTAPAGLIA D	35	CENARGEN	27	7 WILCKE W	96	CIAT
28	13 URQUIAGA S	16	CNPAB	28	7 BALDANI JI	29	CNPAB
29	13 BALDANI JI	28	CNPAB	29	7 SILVA RAMS	66	CNPISA
30	13 DE FREITAS AR	sc	CPPSE	30	7 BUSO JA	97	CNPB
31	13 MARTINEZ ML	sc	CNPGL	31	7 PELOQUIN SJ	57	UWISMAD
32	12 DOS SANTOS HP	229	CNPT	32	7 CABRAL LC	1104	CTAA
33	12 PACHECO CAP	826	CNPMS	33	6 ALVES BJR	46	CNPAB
34	12 PAGLIARINI MS	43	CNPISO	34	6 DE AVILA AC	169	CNPB
35	12 BODDEY RM	53	CNPAB	35	6 GRATTAPAGLIA D	27	CENARGEN
36	12 GONCALVES LC	sc	CNPGL	36	6 LILJENFEIN J	124	UBAYREUT
37	12 DE ALENCAR MM	sc	CPPSE	37	6 ASSIS OBG	60	CNPDIA
38	12 LEAL MDD	sc	CPATC	38	6 LOPES D	126	CTAA
39	12 MADRUGA CR	sc	CNPIC	39	6 MACEDO JLV	351	CPAA
40	11 MORES N	sc	CNPISA	40	6 MIELNICZUK J	131	UFRGS
41	11 AROEIRA LJM	1556	CNPGL	41	5 CURI N	297	UFLA
42	11 VIANA JHM	sc	CNPGL	42	5 JOB AE	334	USP-SC
43	11 KIIHL RAD	1325	CNPISO	43	5 PAGLIARINI MS	34	UEMARINGA
44	11 DE CAMPOS OF	sc	CNPGL	44	5 RAMALHO FS	151	CNPMA
45	11 DE CARVALHO HWL	sc	CPATC	45	5 VIEIRA RF	71	CNPMA
46	10 ALVES BJR	33	CNPAB	46	5 SOUZA GB	163	USP-SC
47	10 EUCLIDES K	sc	CNPIC	47	5 HERRMANN JUNIOR PSP	166	CNPDIA
48	10 EUCLIDES VPB	sc	CNPIC	48	5 VASCONCELLOS CA	92	CNPMS
49	10 FERREIRA AM	sc	CNPGL	49	5 VIDAL PEREZ DANIEL	106	CNPS
50	10 TEODORO RL	sc	CNPGL	50	5 DAVILA AMR	167	FIOCRUZ
51	10 VALLS JFM	299	CENARGEN	51	5 MAGALHAES JR	354	CNPGL
52	10 DE ALMEIDA LA	1304	CNPISO	52	5 PIRES CSS	148	CENARGEN
53	10 BURITY HA	343	EMBRAPA	53	4 BODDEY ROBERT M	35	CNPAB
54	10 BARBOSA PF	sc	CPPSE	54	4 LEHMANN J	1109	UBAYREUT
55	10 MARTIN-NETO L	11	CNPDIA	55	4 ZANUNCIO JC	223	UFV

Note. ClaPa = Classification dans la base Pascal.
ClaWOS = Classification dans la base WOS.
sc = Sans classification

Les colonnes ClaPa, classification dans la base Pascal, et ClaWOS, classification dans la base WOS, indiquent une variation importante entre une base et l'autre. À noter qu'entre les auteurs parmi les dix premiers dans les deux bases, six n'avaient pas de classification (sc) et trois autres parmi les dix premiers auteurs étaient placés en 970^e, 303^e et 205^e position dans l'autre base.

Ce peut s'expliquer par les différences dans les couvertures de journaux scientifiques des deux bases. La base Pascal apparemment n'indexe pas les revues scientifiques de production animale préférées par les chercheurs de l'Embrapa. Ceci est confirmé par l'absence de classification (sc) des chercheurs qui sont entre les 55 premiers dans la WOS et appartiennent à des centres de recherche de production animale, comme Signoretti, R. D., Guidoni, A. L. et Mores, N., de l'Embrapa Porcs et Volailles (CNPSA), Madruga, C. R., Euclides, K. et Euclides, V. P. B. de l'Embrapa Bovins de Viande (CNPGC), de Freitas, A. R., de Alencar, M. M. et Barbosa, P. F., de l'Embrapa Élevage du Bétail Sud-Est (CPPSE), Viana J. H. M., Martinez, M. L., Ferreira, A. M., Goncalves, L. C. Viana, J. H. M., de Campos, O. F. et Teodoro, R. L., de l'Embrapa Bovins de Lait (CNPGL). Les noms de ces auteurs sont marqués en gras dans la colonne à gauche (WOS).

Les journaux scientifiques qui ont reçu les travaux scientifiques de ces chercheurs sont :

Au Brésil : ARQUIVO BRASILEIRO DE MEDICINA VETERINARIA E ZOOTECNIA, REVISTA BRASILEIRA DE ZOOTECNIA-BRAZILIAN JOURNAL OF ANIMAL SCIENCE, PESQUISA AGROPECUARIA BRASILEIRA, BRAZILIAN ARCHIVES OF BIOLOGY AND TECHNOLOGY, PESQUISA VETERINARIA BRASILEIRA e MEMORIAS DO INSTITUTO OSWALDO CRUZ.

À l'extérieur : JAPANESE JOURNAL OF VETERINARY RESEARCH, ANIMAL GENETICS, GENETICS AND MOLECULAR BIOLOGY, BIOLOGY OF REPRODUCTION, VETERINARY PARASITOLOGY, JOURNAL OF VETERINARY MEDICAL SCIENCE e ANIMAL FEED SCIENCE AND TECHNOLOGY.

D'entre eux, seulement le journal "BIOLOGY OF REPRODUCTION" apparaît dans notre échantillon de la base Pascal. De cette façon, c'est une excellente opportunité pour les éditeurs

des revues scientifiques brésiliennes d'amplifier la couverture internationale de leurs journaux et peut-être faire des accords avec l'INIST CNRS France, éditeur de la base Pascal.

La base Pascal

Les 868 références de la base Pascal ont été reformatés et nettoyés sous Infotrans et importés par Dataview qui a généré les matrices année de publication/auteur, année de publication/affiliation et auteur/affiliation pour les années 1998, 1999, 2000, 2001 et 2002. Dataview a aussi généré des matrices symétriques de fréquence d'auteurs pour chaque année de publication qui ont été transformées en cartes de collaboration par le logiciel Matrisme.

Groupes de collaboration

Des matrices symétriques des auteurs de la base Pascal pour les années 1998, 1999, 2000, 2001 et 2002 ont été transformées en cartes de collaboration. Ses figures permettent de visualiser immédiatement les divers groupes d'auteurs qui collaborent et publient ensemble. Le résultat est éclairant. Voici un exemple de quatre groupes :

- a) Mattoso, L. H. C.; Giacometti, J. A.; Job, A. E.; Riul, A. J. R. et Oliveira, O. N. J. R.
- b) Zech, W.; Vilela, L.; Ayarza, M. A.; Neufeldt, H.; Wilcke, W. et Lilienfein, J.
- c) Ascheri, J. L. R.; Cabral, L. C.; Wang, S. H.; Borges, G. G. et Carvalho, C. W. P.
- d) Dobereiner, J.; Baldani, J. I.; Kirchof, G.; Hartmann, A. et Reis, V. M. et Massena Reis, V. (qui sont la même personne).

L'évolution de ces groupes est montrée dans les cartes suivantes :

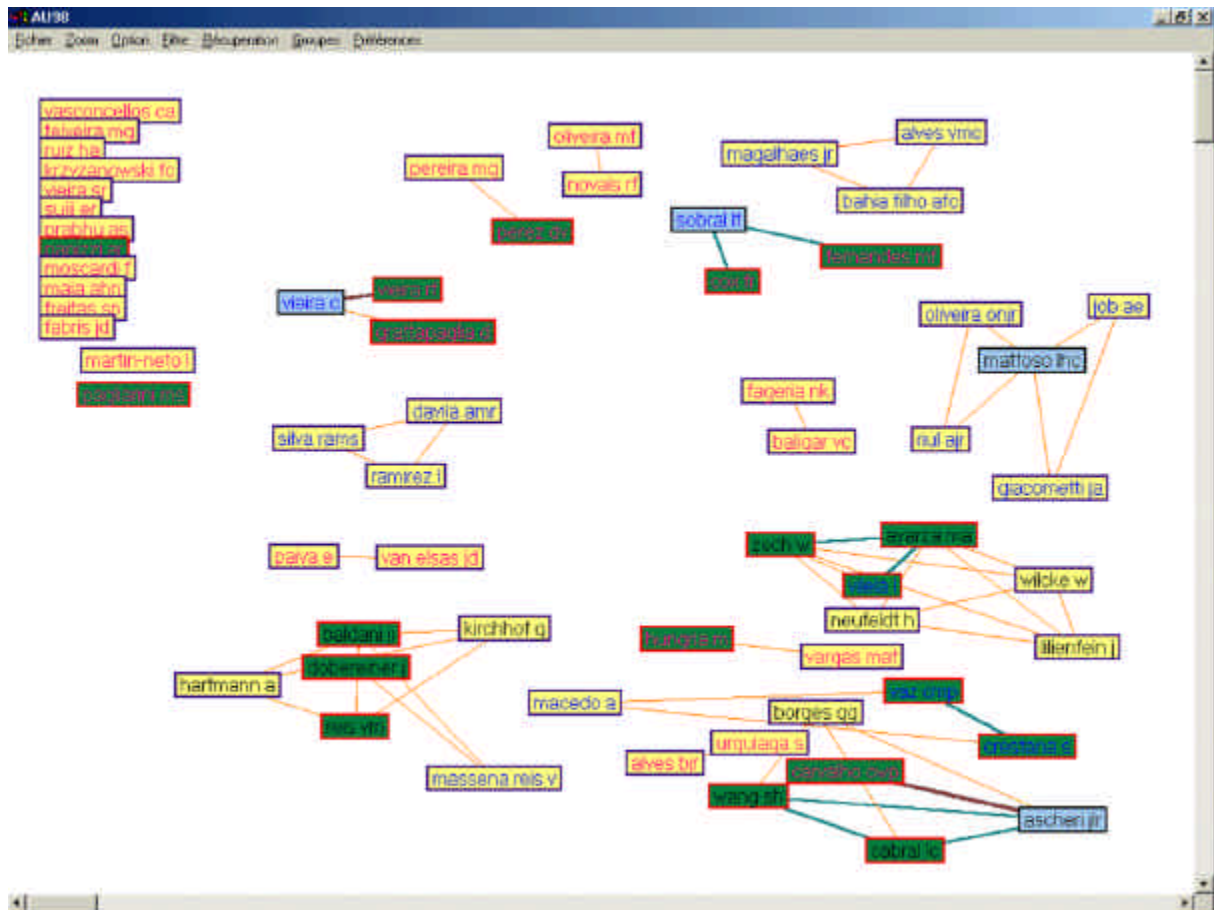


Figure 7 : Carte des groupes d'auteurs, 1998

En 1999, les quatre groupes dont nous avons suivi l'évolution ont connu des sorts différents :

- a) Mattoso, L. H. C.; Giacometti, J. A.; Job, A. E.; Riul, A. J. R. et Oliveira, O. N. J. R. ont gagné Herrmann Junior, P. S. P.
- b) Zech, W.; Vilela, L.; Ayarza, M. A.; Neufeldt, H.; Ont perdu Wilcke, W. et Lilienfein, J. mais ont gagné Resck, D. V. S.; Schroth, G.; Macedo, J. L. V. et Teixeira, W. G.
- c) Ascheri, J. L. R.; Cabral, L. C.; Wang, S. H. et Borges, G. G. ont perdu Carvalho, C. W. P.
- d) Dobereiner, J.; Baldani, J. I. et Kirchhof, G. ont perdu Hartmann, A. et Reis, V. M. et Massena Reis, V. (qui sont la même personne).

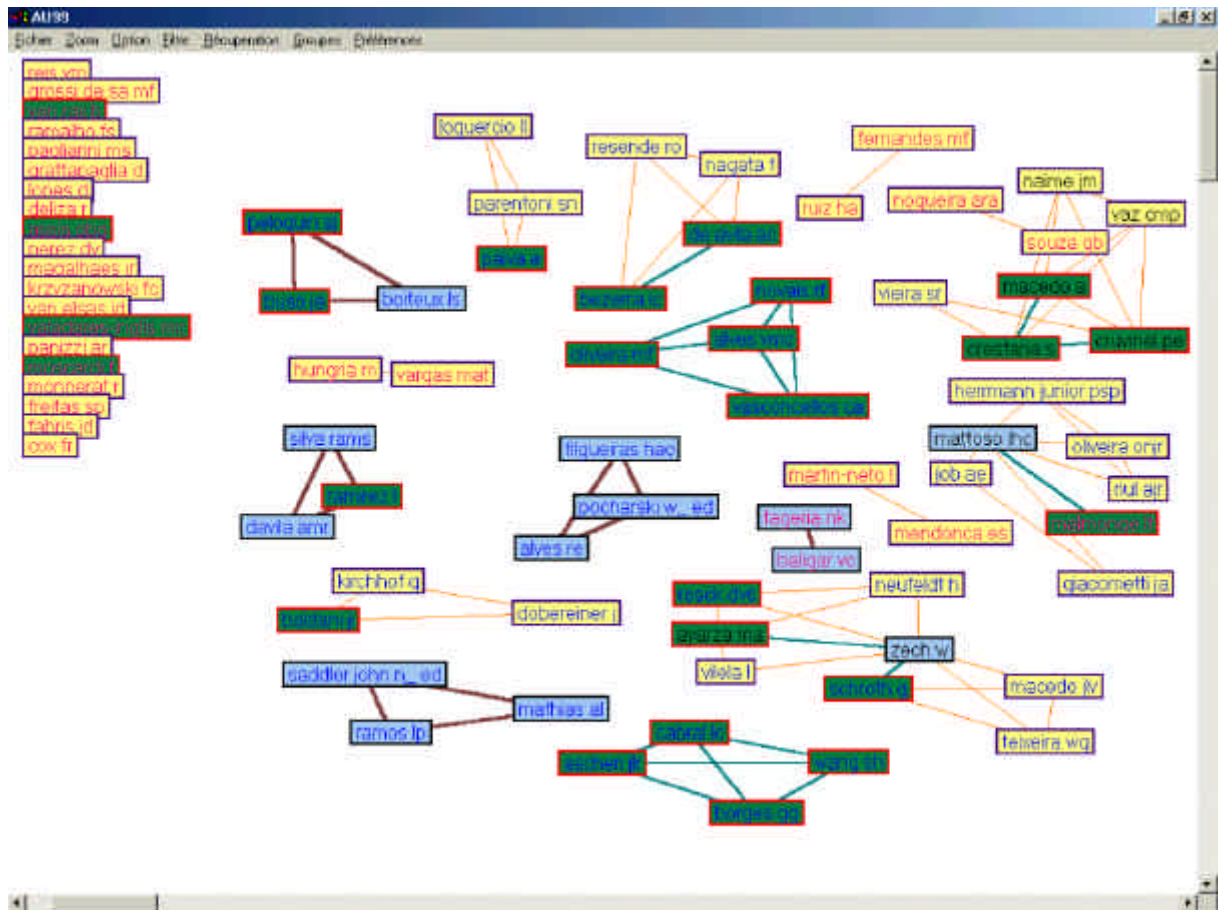


Figure 8 : Carte des groupes d'auteurs, 1999

En 2000, l'évolution des quatre groupes s'est passé comme au-dessous :

- a) Mattoso, L. H. C.; Giacometti, J. A.; Job, A. E. et Herrmann Junior, P. S. P. ont perdu Riul, A. J. R. et Oliveira, O. N. J. R. mais ont gagné Malmonge, L. F.
- b) Zech, W.; Vilela, L.; Ayarza, M. A.; Neufeldt, H.; Schroth, G.; Macedo, J. L. V. et Teixeira, W. G. ont perdu Resck, D. V. S. mais ont gagné de retour Wilcke, W. et Lilienfein, J. et encore Lehmann, J.
- c) Ascheri, J. L. R.; Cabral, L. C.; Wang, S. H. et Borges, G. G. continuent et
- d) Dobereiner, J. et Baldani, J. I. ont perdu Kirchhof, G. mais ont gagné Massena Reis, V.; Urquiaga, S. et Alves, B. J. R.

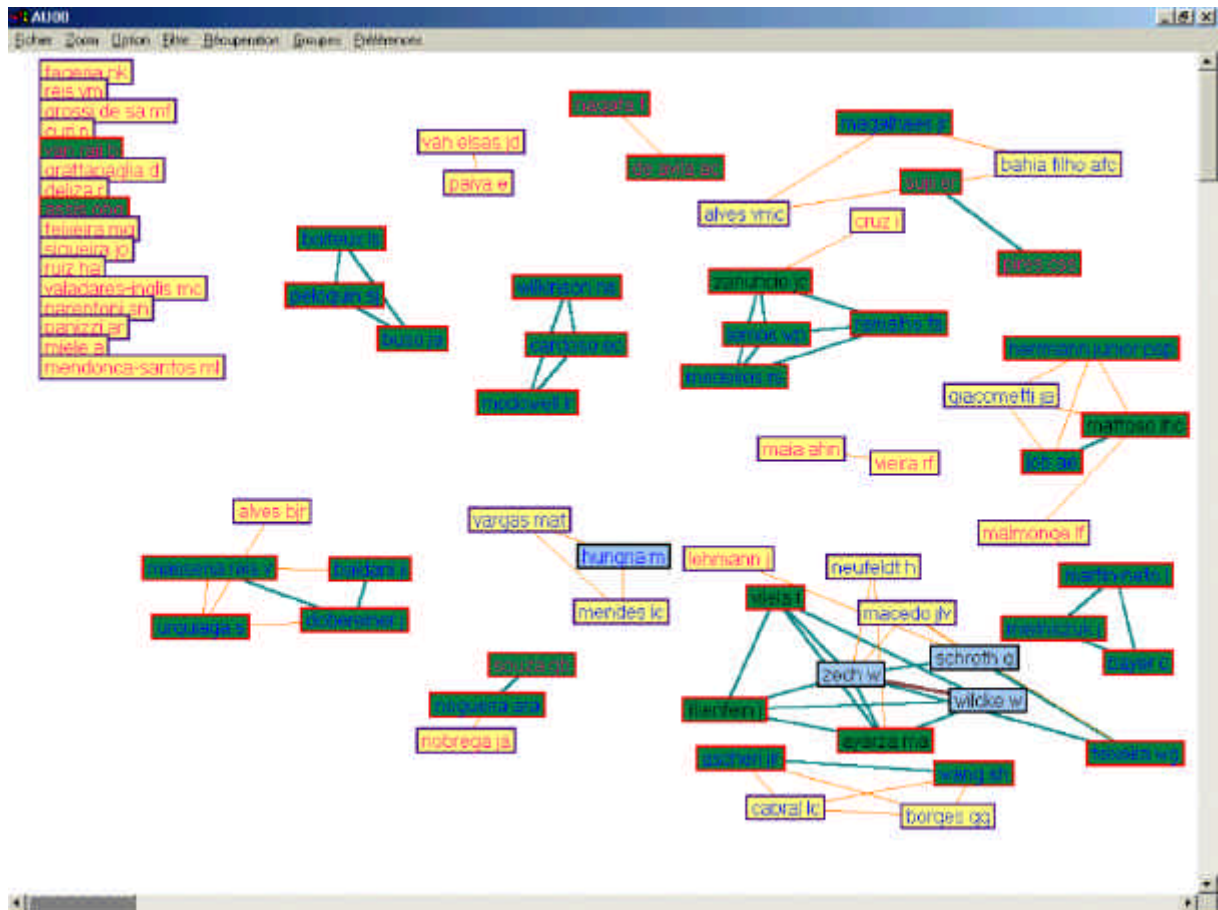


Figure 9 : Carte des groupes d'auteurs, 2000

En 2001, les quatre groupes ont évolué de la façon suivante :

- a) Mattoso, L. H. C.; Job, A. E.; Malmonge, L. F. et Herrmann Junior, P. S. P. ont perdu Giacometti, J. A.
- b) Zech, W.; Vilela, L.; Schroth, G.; Macedo, J. L. V.; Wilcke, W.; Lilienfein, J. et Lehmann, J. ont perdu Ayarza, M. A.; Neufeldt, H. et Teixeira, W. G.
- c) Ascheri, J. L. R.; Cabral, L. C. et Wang, S. H. ont perdu Borges, G. G. et il y a eu un réarrangement au dernier groupe
- d) Urquiaga, S.; Reis, V. M. et Alves, B. J. R. ont perdu Dobereiner, J. et Baldani, J. I. mais ont gagné Kirchhof, G.; Hartmann, A. et Boddey Robert M.

Un autre groupe bien nombreux est apparu :

- e) Vaz, C. M. P.; Martin Neto, L.; Crestana, S.; Bayer, C.; Mendonca E. S.; Mielniczuk, J.; Naime, J. M.; Macedo, A.; Ferreira, M. M.; Curi, N. et Silva, M. L. N.

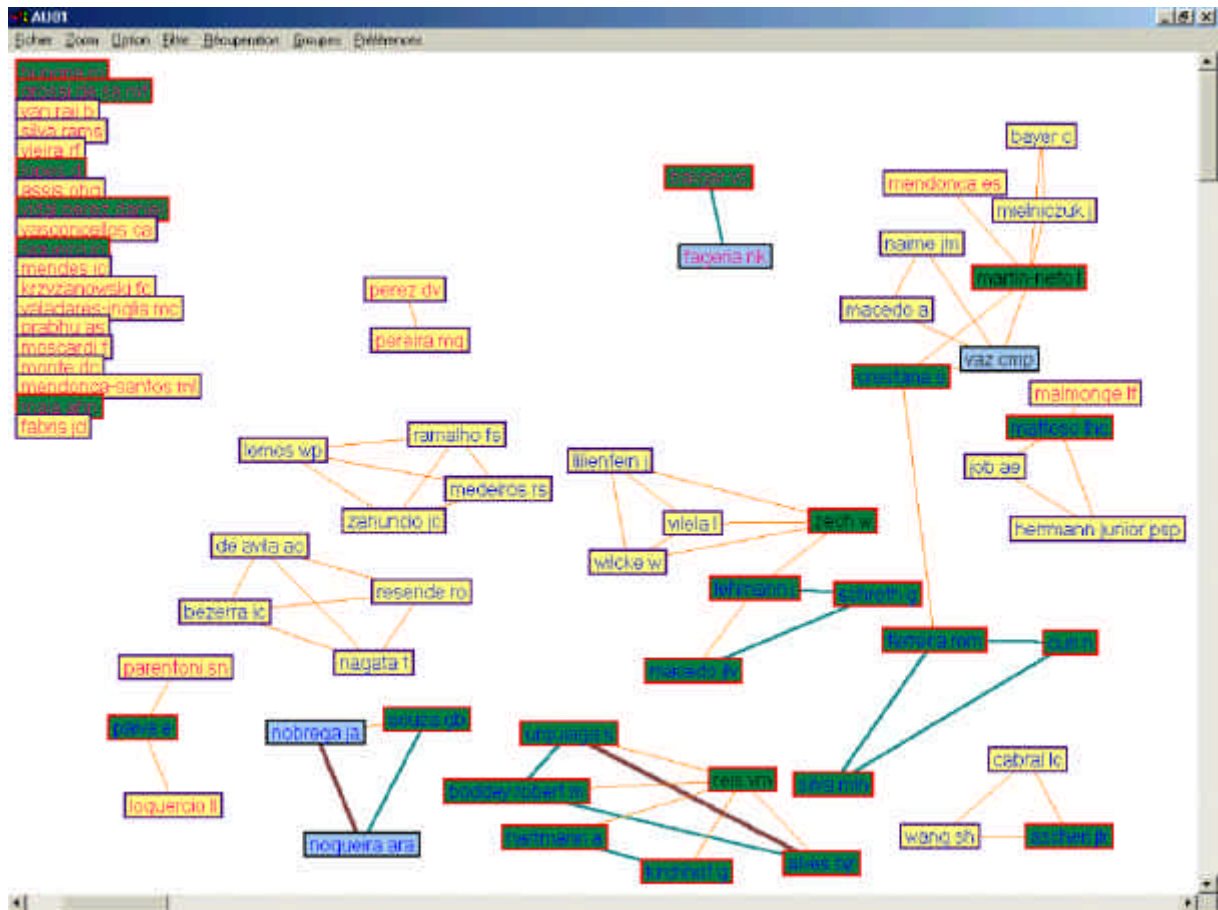


Figure 10 : Carte des groupes d'auteurs, 2001

Finally, in 2002, the five groups that we are following now have known several rearrangements. The large group that emerged in 2001 has almost absorbed the first group :

- a) Vaz, C. M. P.; Martin Neto, L.; Crestana, S.; Mielniczuk, J.; Naime, J. M.; Macedo, A.; ont perdu Bayer, C.; Mendonca E. S.; Ferreira, M. M.; Curi, N. et Silva, M. L. N. mais ont gagné Herrmann Junior, P. S. P.

The second group has split into two and has lost Zech, W.; Vilela, L.; Wilcke, W.; Lilienfein, J.; Lehmann, J.

- b) Ayarza, M. A.; Neufeldt, H. et Resck, D. V. S. et
- c) Schroth, G.; Macedo, J. L. V.

And the other groups have also been modified :

- d) Ascheri, J. L. R. et Wang, S. H. ont perdu Cabral, L. C. et ont gagné Carvalho C. W. P. et Oliveira, M. F.
- e) Urquiaga, S. et Reis, V. M. ont perdu Alves, B. J. R.; Kirchhof, G.; Hartmann, A. et Boddey Robert M. mais ont gagné Dobereiner, J. et Baldani, J. I.

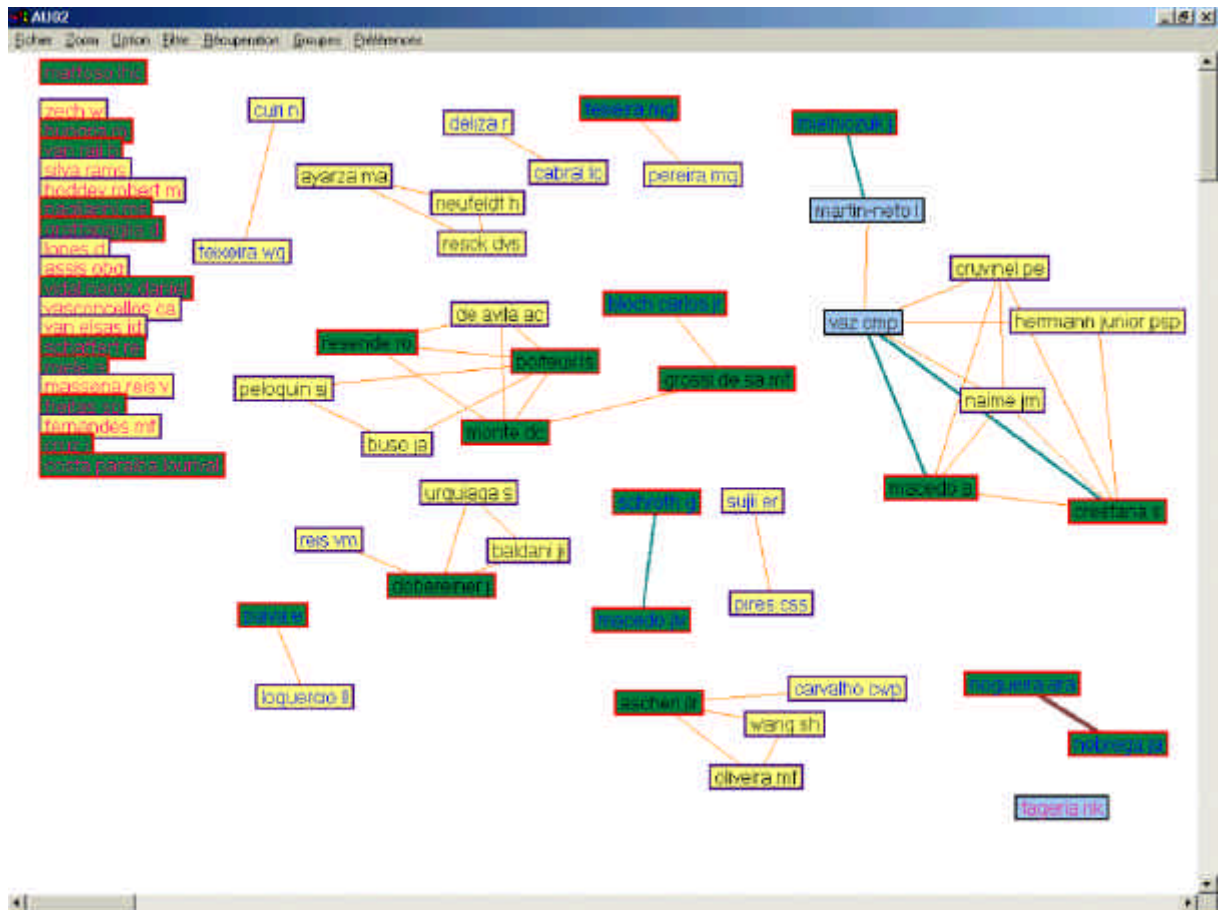


Figure 11 : Carte des groupes d'auteurs, 2002

L'analyse des institutions

Les 868 références de la base Pascal avec Embrapa comme affiliation avaient 649 références des années 1998, 1999, 2000, 2001 et 2002. Ces 649 références réunissaient des citations de 1903 auteurs et 1608 institutions (noms et adresses). Les institutions d'affiliation ont été listées par ordre alphabétique et, ensuite, les noms ont été normalisés. Le critère adopté a été celui de déconsidérer les institutions avec une seule citation et privilégier les institutions avec au moins deux participations. Furent identifiés 253 institutions, divisées en six groupes : Embrapa (35), Brésil (66), Europe (70), EUA/Canada (47), Afrique/Amérique Latine/Asie/Océanie (30) et Système ONU (5). Nous avons préservé les institutions avec fréquence unitaire en les réunissant dans les rubriques Autres Brésil, Autres Europe, Autres EUA/Canada et suivantes.

Le groupe Embrapa

Le groupe Embrapa a été isolé pour un examen en profondeur. Les quatre centres de recherche qui ont le plus de participations en travaux publiés dans la base Pascal sont: Embrapa Ressources Génétiques et Biotechnologie (Cenargen) ; Embrapa Instrumentation Agricole (CNPDI) ; Embrapa Agro-industrie d'Aliments (CTAA) et Embrapa Agrobiologie (CNPAB). Quatre centres de recherche, Embrapa Surveillance par Satellite (CNPM), Embrapa Café (CAFE), Embrapa Amapa (CPAF-AP) et Embrapa Roraima (CPAF-RR) n'ont pas de travaux publiés entre 1998/2002 dans la base Pascal. Le tableau 29 présente les statistiques du groupe Embrapa. A remarquer que 12 centres publient moins d'un article par année dans les revues indexées par la base Pascal.

Tableau 29 : Base Pascal ; Groupe Embrapa

Groupe Embrapa	2002	2001	2000	1999	1998	Total	Part/an
1 CENARGEN	15	10	25	18	8	76	15,2
2 CNPDIA	17	17	10	14	10	68	13,6
3 CTAA	9	9	9	7	12	46	9,2
4 CNPAB	10	10	8	4	9	41	8,2
5 CNPSO	3	8	8	7	10	36	7,2
6 CNPMA	7	7	8	7	7	36	7,2
7 CPAC	5	5	8	13	7	38	7,6
8 CNPMS	7	7	5	8	7	34	7
9 CNPAF	8	16	1	6	4	35	6,8
10 CPAA	7	7	8	3	1	26	5,2
11 CNPS	8	7	2	4	4	25	5
12 CNPF	4		1	17	2	24	4,8
13 CNPH	3	2	10	8		23	4,6
14 CNPMF	4	4	6	3	4	21	4,2
15 CPPSE	5	6	3	2	3	19	3,8
16 CPATU	4	4	5	2	3	18	3,8
17 CNPAT	4	1	2	7	2	16	3,2
18 CNPSA	3	2	1	6	2	14	2,8
19 CPACT	3	6	2	1	1	13	2,6
20 CNPT		3	4	1	5	13	2,6
21 CNPGC	3	2	4	1		10	2
22 CPAP		1	3	3	2	9	1,8
23 CNPA		1	5	1	1	8	1,6
24 CNPUV	2	2	1	2		7	1,4
25 CPATC	2			1	4	7	1,4
26 CPATSA	3	1	1	1	1	7	1,4
27 CNPGL	1		3	3		7	1,4
28 SIÈGE	1			1	2	4	0,8
29 CPAF-AC		1	1	1		3	0,6
30 CNPTIA		2			1	3	0,6
31 CPAO		1		1		2	0,4
32 CPAF-RO	1					1	0,2
33 CPAMN			1			1	0,2
34 CNPC	1					1	0,2
35 CPPSUL					1	1	0,2
36 CNPM						0	0
37 CPAF-AP						0	0
38 CAFE						0	0
39 CPAF-RR						0	0
Total	140	143	145	153	113	694	138,8

Ensuite, ont été produites les matrices auteur/affiliation (groupes), auteur/revues scientifiques, auteur/langue, groupes d'affiliation/année de publication, groupe Embrapa/autres groupes, groupe Embrapa/revues scientifiques et groupe Embrapa/langue. Même si les statistiques avec les auteurs sont aussi éclairantes, nous avons opté pour étudier les matrices affiliation (groupes) et Embrapa, en ce qui concerne les centres de recherche.

L'évolution des partenariats

La matrice institutions d'affiliation/année de publication permet de regarder l'évolution des partenariats au long de la période étudiée. Les institutions que ne sont pas Embrapa et qui apparaissent entre les 40 premiers en participation dans articles publiés (avec fréquence jusqu'à six) sont Université Fédérale de Viçosa, UFV (50) ; Université Fédérale du Paraná, UFPR (43) ; Université de Sao Paulo, Sao Carlos, USP-SC (37) ; Université Fédérale du Rio de Janeiro, UFRJ (34) ; Université Fédérale Rurale du Rio de Janeiro, UFRRJ (33) ; **Agricultural Research Service**, USDA-ARS (28) ; **Escola Superior de Agricultura Luiz de Queiroz**, ESALQ (25) et le Centre de Coopération Internationale en Recherche Agronomique pour le Développement, CIRAD (25). Ces et les autres institutions non-Embrapa sont marquées en gris au tableau 30.

Une autre observation est qu'il y a une régularité dans l'activité de publication de la majorité des centres de recherche et d'institutions. A remarquer trois institutions ; Institut Brésilien de l'Environnement et des Ressources Naturelles Renouvelables (IBAMA), Université de British Columbia (UBRITCOLUMBIA) et Agence Internationale d'Energie (AIENERGIE) qui ont participé d'un nombre élevé d'articles (15) dans une seule année, 1999. Dans la majorité des cas, les collaborations génèrent des publications tous les ans, ce qui indique une relation profitable pour les deux parties. Ces détails sont fournis tableau 30.

Tableau 30 : Base Pascal ; Participation en articles par institution et année

Partenaires non-Embrapa	2002	2001	2000	1999	1998	Total	Part/an
1 UFV	6	10	12	8	14	50	10
2 UFPR	2		3	37	1	43	8,6
3 USP-SC	7	7	8	9	6	37	7,4
4 UFRJ	9	8	2	8	7	34	6,8
5 UFRRJ	8	7	5	7	6	33	6,6
6 USDA-ARS		6	7	10	5	28	5,6
7 ESALQ	7	4	3	6	5	25	5
8 CIRAD	9	6	1	6	3	25	5
9 UFSCAR	7	6	2	8		23	4,6
10 UNICAMP	5	3	4	6	4	22	4,4
11 UFLORIDA		3	13	2	4	22	4,4
12 UNB	7	5	1	6	2	21	4,2
13 UBAYREU	2	3	10	4	2	21	4,2
14 UFLA	2	11	2	2	4	21	4,2
15 UNESP	4	3	6	4	1	18	3,6
16 UCORNEL		4	5	5	4	18	3,6
17 CIAT	1	2	6	3	5	17	3,4
18 UCAL	3	8	4		2	17	3,4
19 IAC-APTA	4		2	5	5	16	3,2
20 UWAGENINGEN	2	4	4	4	2	16	3,2
21 AIENERGIE				15		15	3
22 UBRITCOLUMBIA				15		15	3
23 IBAMA				15		15	3
24 UFPEL	3	6		5		14	2,8
25 UFMG	1	3	5	3	1	13	2,6
26 UEL	3	4	2	1	2	12	2,4
27 USP	3	2	2	4	1	12	2,4
28 UFRGS	5	2	2	2	1	12	2,4
29 UWISCMAD	2		5	4		11	2,2
30 UMARINGA	2	2	1	2	3	10	2
31 UHAMBURG	4	1	4			9	1,8
32 INPA	2	4	1	2		9	1,8
33 CENA	2		4		2	8	1,6
34 INRA	3	1	2	1	1	8	1,6
35 CNPQ		1	3	3		7	1,4
36 NCSU	1			1	5	7	1,4
37 ZEF	3	2	1	1		7	1,4
38 UFU		2	4			6	1,2
39 UFPB	3		2	1		6	1,2
40 UGOETTINGEN	3		1	1	1	6	1,2
41 FIOCRUZ	1			5		6	1,2
Total 40 premières	126	130	139	221	99	715	143
Total participations	169	180	189	260	136	934	186,8

Note. Part/an = Participation en articles par année.

Ensemble, les chercheurs du groupe étudié de 253 institutions ont généré 1743 participations dans les 649 articles publiés, en moyenne, 348 participations par an entre 1998 et 2002, justifiant ainsi la mention du groupe Embrapa et ses collaborateurs entre les 15 institutions qui publient le plus grand nombre d'articles scientifiques au Brésil²⁰. Ces chiffres sont montrés tableau 31.

Tableau 31 : Base Pascal ; Participation en articles par région géographique et année

	2002	2001	2000	1999	1998	Total	Part/an
Embrapa	140	143	145	153	113	694	138,8
Brésil	113	111	97	168	89	578	115,6
Europe	49	50	44	35	20	198	39,6
EUA/Canada	21	28	57	56	27	189	37,8
Système ONU	2	3	7	18	6	36	7,2
Afrique/Amérique Latine/Asie	10	11	6	12	9	48	9,6
Total de participations en articles	335	346	356	442	264	1743	348,6

Note. Part/an = Participation en articles par année.

La vision des partenariats par centres de recherche

La matrice groupe Embrapa/année de publication permet d'évaluer l'action d'une unité de recherche, en détectant par exemple, avec quelles institutions elle travaille et, suivant la ligne du temps, comment ces collaborations évoluent. Nous avons choisi d'étudier les deux premiers centres. Les partenaires les plus significatifs pour l'Embrapa Ressources Génétiques et Biotechnologie (Cenargen) sont l'Université de Brasilia (UNB) et l'Université Fédérale de Viçosa (UFV). Pour l'Embrapa Instrumentation Agricole (CNPDIA) le portrait change. Les partenaires principaux sont l'Université de Sao Paulo à Sao Carlos (USP-SC), l'Université Fédérale de Sao Carlos (UFSCAR), l'Université Paulista (UNESP) et l'Université de Californie (UCAL). En mettant les partenariats des deux centres en perspective il est possible de dénoter que même en participant moins fois dans des articles, 68 pour 76 occurrences

²⁰ Voyez plus de détails dans le chapitre "Indicateurs de R&D au Brésil".

(tableau 29), le CNPDIA englobe un nombre plus grand de partenaires en ses articles que le Cenargen, 138 pour 88 occurrences. Ces collaborations sont montrés dans le tableau 32.

Tableau 32 : Base Pascal ; Partenaires de l'Embrapa Ressources Génétiques et Biotechnologie (Cenargen), à gauche et de l'Embrapa Instrumentation Agricole (CNPDI A), à droite

Cenargen	2002	2001	2000	1999	1998	Total	CNPDI A	2002	2001	2000	1999	1998	Total
UNB	5	1	1	2	2	11	USP-SC	7	6	7	8	6	34
UFV			3	1	3	7	UFSCAR	3	2	1	7		13
UPURDUE	1		1		1	3	UNESP		1	3	3		7
NORTHERNARIZUN			3			3	UCAL	2	3			2	7
UCB	2	1				3	UFRGS	3	1	2			6
CIRAD			1	1	1	3	UFV	1	2	1	2		6
UWAGENNIGEN	1	1	1			3	UFPE	1	1	1		1	4
USDA-ARS			1	2		3	UESC	1	1	2			4
UCAL			2			2	UFMG			4			4
SAINSBURYLAB-UK		1		1		2	UFOP			4			4
IAC-APTA				1	1	2	CENA	2					2
UFPR			1	1		2	USP		1	1			2
UCORNEL			1	1		2	IAC-APTA	1			1		2
UNESP	1		1			2	UFMS			2			2
UGEORGIA	1		1			2	MAHATGANDHIUNI	2					2
UFRN	1		1			2	CNPQ			2			2
UFLORIDA				2		2	UFRJ		1		1		2
IACR-UK			1	1		2	UPENNS			1	1		2
UATHENS	1	1				2	Total 1	23	19	31	23	9	105
UFRRJ				2		2	Autres EUA+Canada	4	2		2		8
Total 1	13	5	19	15	8	60	Autres Europe	4	2		1	2	9
Autres Europe	4	4	5	2		15	Autres Brésil	4	4	1			9
Autres EUA+Canada	1		2	3		6	Autres AFR/AMLAT/ASIE	3				1	4
Autres Brésil	2		2	1		5	Autres Oceanie		2			1	3
Autres AFR/AMLAT/ASIE			1		1	2	Total 2	38	29	32	26	13	138
Total 2	20	9	29	21	9	88							

L'étude des revues scientifiques

La matrice groupe Embrapa/revues scientifiques identifie la politique de publication des centres de recherche. Voyez, par exemple, dans le tableau 33, respectivement, ces analyses pour l'Embrapa Ressources Génétiques et Biotechnologie (Cenargen), à gauche et l'Embrapa Instrumentation Agricole (CNPDI A), à droite. Les journaux préférés des chercheurs de

l'Embrapa Ressources Génétiques et Biotechnologie sont : Euphytica, des Pays Bas ; Environmental Entomology, des États Unis; Theoretical and Applied Genetics, d'Allemagne ; et Archives of Virology, de l'Autriche. Tandis que les chercheurs de l'Embrapa Instrumentation Agricole préfèrent : Revista Brasileira de Ciência do Solo, du Brésil ; Polymer, du Royaume-Uni ; Soil & Tillage Research, des Pays Bas ; Journal of Applied Polymer Science, des États Unis et Synthetic Metals, de la Suisse.

Tableau 33 : Base Pascal ; Revues scientifiques préférées à l'Embrapa Ressources Génétiques et Biotechnologie, à gauche et à l'Embrapa Instrumentation Agricole, à droite

Revue scientifique Cenargen		Revue scientifique CNPDIA	
Euphytica (Pays-Bas)	10	Revista brasileira de ciencia do solo (Brésil)	6
Environmental entomology (États-Unis)	4	Polymer : (Guildford) (Royaume-Uni)	5
Theoretical and Applied Genetics (Allemagne)	4	Soil & tillage research (Pays-Bas)	5
Archives of virology (Autriche)	3	Journal of applied polymer science (États-Unis)	4
FEMS microbiology letters (Pays-Bas)	2	Soil Science Society of America journal (États-Unis)	4
Fruits : (1978) (France)	2	Synthetic metals (Suisse)	4
Journal of applied entomology : (1986) (Allemagne)	2	Geoderma : (Amsterdam) (Pays-Bas)	3
Journal of chemical ecology (États-Unis)	2	Journal of non-crystalline solids (Pays-Bas)	3
Journal of chromatography (Pays-Bas)	2	Computers and electronics in agriculture (Pays-Bas)	2
Journal of Experimental Botany (Royaume-Uni)	2	Electrochimica acta (Royaume-Uni)	2
Journal of invertebrate pathology : (Print) (États-Unis)	2	Journal of environmental quality (États-Unis)	2
Journal of plant nutrition (États-Unis)	2	Powder technology (Suisse)	2
Nematology : (Leiden) (Pays-Bas)	2	Soil science (États-Unis)	2
Toxicon : (Oxford) (Royaume-Uni)	2		

L'étude des langues

La matrice groupe Embrapa/langue permet de vérifier la multi-culturalité des centres de recherche. La majorité écrasante, 82% des participations en articles scientifiques réunis dans cet échantillon de la base Pascal est de langue anglaise. En ordonnant la matrice par la fréquence décroissante en portugais, nous vérifions que quelques centres comme l'Embrapa Semi-Aride (CPATSA), l'Embrapa Rondonia (CPAF-RO) et l'Embrapa Agriculture et Élevage Ouest (CPAO) ont la majorité de leur production scientifique dans la langue nationale. Ils sont marqués en gris dans le tableau 34.

Tableau 34 : Base Pascal ; Langues des articles

	Anglais	Portugais	Espagnol	Français	Allemand	Italien	Total
CNPS	13	11	1				25
CNPMS	26	8					34
CNPMF	11	7	2	1			21
CNPDIA	62	6					68
CNPAF	27	6		2			35
CNPT	8	5					13
CPAC	29	4	1	4			38
CPATSA	3	4					7
CNPSO	33	3					36
CNPMA	33	3					36
CPACT	9	3			1		13
CNPAB	39	2					41
CPAA	23	2	1				26
CPATC	3	2	2				7
CPAO		2					2
CENARGEN	75	1					76
CPPSE	18	1					19
CPATU	16	1		2			19
CNPGC	9	1					10
CPAF-AC	2	1					3
CNPTIA	2	1					3
CPAF-RO		1					1
CTAA	23		22			1	46
CNPF	24						24
CNPH	23						23
CNPAT	14		1	1			16
CNPSA	13			1			14
CPAP	9						9
CNPA	7		1				8
CNPUV	3			4			7
CNPGL	7						7
SIÈGE	4						4
CPPSUL	1						1
CNPC	1						1
CPAMN	1						1
Total	571	75	31	15	1	1	694

La base WOS

Les 2.421 références de la base WOS avaient 1.627 références des années 1998, 1999, 2000, 2001 et 2002. Ces 1.627 références réunissaient des citations de 3.521 auteurs et 2.239 institutions (noms et adresses). Ces noms et adresses ont été listés par ordre alphabétique et, ensuite, normalisés. Le critère adopté a été le même de celui choisi pour la base Pascal : Celui de privilégier les institutions avec, au moins, deux citations. Ont été identifiés 327 institutions divisées en six groupes : Embrapa (39), Brésil (110), Europe (80), EUA/Canada (59), Afrique/Amérique Latine/Asie (36) et Système ONU (3). Nous avons préservé les institutions avec fréquence unitaire en les réunissant dans les rubriques Autres Brésil, Autres Europe, Autres EUA/Canada et suivantes.

Le groupe Embrapa

Le groupe Embrapa a été isolé pour être examiné davantage. Il est présenté dans le tableau 35. On observe que les 39 centres de recherche de l'Embrapa sont présents. Trois centres ont publié moins d'un article par an pendant la période étudiée.

Tableau 35 : Base WOS ; Groupe Embrapa

Groupe Embrapa	2002	2001	2000	1999	1998	Total	Part/an
1 CENARGEN	61	43	59	41	23	227	45,4
2 CNPGL	20	20	21	29	17	107	21,4
3 CNPSO	15	14	21	12	19	81	16,2
4 CNPAF	20	21	13	12	15	81	16,2
5 CNPDIA	17	18	14	14	12	75	15
6 CPPSE	17	14	13	17	13	74	14,8
7 CNPMA	16	9	15	21	12	73	14,6
8 CPAC	15	7	22	18	11	73	14,6
9 CNPMS	16	10	16	17	11	70	14
10 CNPSA	7	8	16	18	20	69	13,8
11 CNPAB	11	15	16	11	12	65	13
12 CNPGC	16	11	18	11	10	66	13,2
13 CPATSA	8	5	8	21	13	55	11
14 CPAP	10	9	10	8	15	52	10,4
15 CPATU	11	7	13	7	6	44	8,8
16 CNPT	7	4	13	7	9	40	8
17 CNPH	5	6	10	13	4	38	7,6
18 CNPS	12	4	7	8	5	36	7,2
19 CPAA	11	7	11	4	2	35	7
20 CNPAT	11	2	9	7	4	33	6,6
21 CTAA	6	5	9	5	4	29	5,8
22 CNPC	3	1	3	13	8	28	5,6
23 CPATC	5	4	8	4	3	24	4,8
24 CNPMF	5	4	5	6	3	23	4,6
25 CNPA	2	3	9	5	2	21	4,2
26 SIÈGE	8	2		6	3	19	3,8
27 CPACT	1	6	5	2	3	17	3,4
28 CNPF	7	2	4		3	16	3,2
29 CNPUV	1	2	3	7	1	14	2,8
30 CPAO	1	1	5	6	1	14	2,8
31 CPPSUL	5		1	1	1	8	1,6
32 CPAF-AC		2	1	3	1	7	1,4
33 CPAF-RO	2	1		1	2	6	1,2
34 CNPTIA	2	1			2	5	1
35 CPAF-AP			4	1		5	1
36 CPAMN	1	2	1		1	5	1
37 CPAF-RR	1		1			2	0,4
38 CNPM					1	1	0,2
39 CAFE	1					1	0,2
Total	357	270	384	356	272	1639	327,8

Note. Part/an = Participation en articles par année.

Les cinq centres de recherche qui ont le plus de participations dans des travaux publiés dans la base WOS sont : Embrapa Ressources Génétiques et Biotechnologie, CENARGEN (227) ; Embrapa Bovins de Lait, CNPGL (107) ; Embrapa Riz et Harricot, CNPAF (81) ; Embrapa Soja, CNPSO (81) et Embrapa Instrumentation Agricole, CNPDIA (75).

Les deux unités qui apparaissaient en troisième et quatrième places dans la base Pascal, Embrapa Agro-industrie d'Aliments (CTAA) et Embrapa Agrobiologie (CNPAB), apparaissent maintenant, respectivement, en vingt et unième et onzième places. Autre centre à mentionner en fonction de son score dans les bases Embrapa - Production scientifique et Collection documentaire - Embrapa Maïs et Sorgho (CNPMS) apparaît en neuvième place dans les deux bases, Pascal et WOS.

Les caractéristiques des distributions par régions de l'Embrapa

Les participations distribuées par régions géographiques auxquelles appartiennent les centres de recherche de l'Embrapa est différente de celle enregistré par Gregolin et al. (2002 et 2005)²¹ où prédominent, avec un grand marge les institutions de la région Sud-Est (77%) suivi du Sud (15%), Nord-Est (9%), Centre-Ouest (4%) e Nord (2%). À l'Embrapa cette balance scientifique est plus équilibrée. La région Centre-Ouest occupe la première place (35%), suivant les régions Sud-Est (33%), Sud (15%), Nord-Est (9%) et Nord (8%). Cette information a été croisée avec les années étudiées pour accompagner l'évolution dans la période, montrée tableau 36.

Tableau 36 : Base WOS ; Participations en articles de l'Embrapa par région et année

Embrapa par région	2002	2001	2000	1999	1998	Total	%
Centre-Ouest	137	100	137	115	82	571	35
Sud-Est	117	96	111	122	89	535	33
Sud	43	36	63	47	56	245	15
Nord-Est	24	19	34	49	30	156	9
Nord	36	19	39	23	15	132	8
Total	357	270	384	356	272	1639	100

Entre les centres de recherche de l'Embrapa, ces des régions Sud et Nord-Est ont réduit leur participations dans les articles publiés dans la période. Dans les centres du Sud la participation a décru de 34% à partir de 2000 et dans ces du Nord-Est, de 50% depuis 1999. Dans ces du Centre-Ouest il y a eu une croissance de 74% suivie par le Sud-Est avec 27% et le Nord, de 157%. Cette dernière région a enregistré une réduction du nombre de participation en articles entre 2000 et 2002 mais le résultat final est plus du double de 1998.

En adoptant les mêmes critères établis pour la base Pascal, ont été produites des matrices groupes d'affiliation/année de publication, groupe Embrapa/autres groupes, groupe Embrapa/revues scientifiques et groupe Embrapa/langue.

Les partenaires du groupe Brésil

Cette fois, nous avons choisi d'étudier la matrice groupe Embrapa/autres groupes par le critère des institutions non-Embrapa. La question formulée a été : Quels sont les partenaires des centres de recherche de l'Embrapa? Dans la base WOS, les trois institutions qui publient le plus en partenariat avec Embrapa sont l'Université Fédérale de Viçosa, UFV (127 participations), l'Université de Brasilia, UNB (78 participations), l'Université Paulista, UNESP (58 participations) et l'Université Fédérale de Minas Gerais, UFMG (52 participations). A remarquer que l'Université de Brasilia (UNB) concentre 90% de ses participations avec les centres de Brasilia, marqués en gris, Embrapa Ressources Génétiques et Biotechnologie (Cenargen), Embrapa Cerrados (CPAC), Embrapa Horticulture (CNPQ) et le Siège d'Embrapa, un total de 70 en 78 participations. Les listes figurent au tableau 37.

²¹ Voyez plus de détails dans le chapitre "Indicateurs de R&D au Brésil".

Tableau 37 : Base WOS ; Partenaires Embrapa de l'UFV, l'UNB, l'UNESP et l'UFMG

UFV		UNB		UNESP		UFMG	
CNPGL	33	CENARGEN	56	CNPGL	10	CNPMS	19
CPATSA	26	CPAC	8	CNPGC	9	CNPGL	13
CNPSA	12	CPAP	4	CNPDIA	7	CENARGEN	7
CNPMS	12	CNPH	4	CENARGEN	6	CPAC	4
CENARGEN	10	SIÈGE	2	CPPSE	6	CNPSA	2
CNPGC	6	CNPSO	1	CPAP	5	CNPC	2
CNPA	5	CNPAF	1	SIÈGE	3	CPAP	2
CNPMA	4	CPATU	1	SIÈGE	3	CNPUV	1
CNPSO	2	CNPMF	1	CPAC	3	SIÈGE	1
CPATC	2	Total	78	CPAMN	2	CNPGC	1
CNPAB	2			CNPAB	1	Total	52
SIÈGE	2			CPPSUL	1		
CNPC	2			CPATC	1		
CNPDIA	2			CNPMA	1		
CPPSE	2			CPATU	1		
CPAF-RO	1			CTAA	1		
CNPUV	1			CNPMS	1		
CPATU	1			Total	58		
CTAA	1						
CNPH	1						
Total	127						

Les caractéristiques régionales de la coopération au Brésil

Quand le groupe de partenaires au Brésil est décomposé par régions réapparaît la tendance de super valorisation de la région Sud-Est commune dans les statistiques brésiliennes de production scientifique (Gregolin et d'autres, 2002 et 2005). Par contre, la région Centre-

Ouest passe de la quatrième à la deuxième place, ensemble avec la région Sud. Les valeurs du tableau 39 en témoignent.

Tableau 39 : Base WOS; Partenaires Embrapa au Brésil par région et année

Brésil par région	2002	2001	2000	1999	1998	Total	%
Sud-Est	175	120	126	144	98	663	62
Centre-Ouest	45	36	25	27	20	153	14
Sud	41	45	40	11	17	154	14
Nord-Est	25	20	8	15	21	89	8
Nord	9	4	1	2	2	18	2
Total	295	225	200	199	158	1077	100

Les partenariats dans le monde

Nous avons cherché ensuite les centres de recherche qui ont des caractéristiques le plus internationales, en travaillant avec des partenaires de toutes les quatre groupes géographiques créés, EUA/Canada, Europe, Afrique/Amérique Latine/Asie et Système ONU. Ils sont identifiés en gris dans le tableau 38.

Tableau 38 : Base WOS ; Les centres de l'Embrapa les plus internationaux.

EUA/Canada		Europe		AFR/AMLAT/ASIE		Sys ONU	
CENARGEN	44	CENARGEN	41	CENARGEN	10	CPAC	14
CNPAF	16	CPAC	20	CNPAB	5	CNPAF	2
CNPH	16	CNPAB	17	CPATU	5	CENARGEN	1
CPATU	11	CPAA	16	CNPGC	5	SIÈGE	1
CNPMS	10	CNPMA	10	CNPMA	4	CNPDIA	1
CNPDIA	9	CPATU	9	CNPSO	4	CPATU	1
CPAC	9	CNPSO	6	CNPDIA	3	CNPGC	1
CNPSO	8	CNPH	6	CPAC	3	CNPMF	1
CPAP	7	CNPS	5	CPPSUL	2	CNPMS	1
CNPSA	5	CNPAT	5	CNPSA	1	Total	23
CNPGL	5	CNPAF	3	CNPAF	1		
CNPMA	4	CNPT	3	CNPT	1		
SIÈGE	4	CPACT	3	SIÈGE	1		
CNPC	4	CNPGC	3	CPACT	1		
CNPGC	4	CNPF	3	CTAA	1		
CNPA	4	CNPMS	3	CNPH	1		
CNPAB	3	CNPSA	2	Total	48		
CPATSA	3	CNPUV	2				
CPAA	3	CPATSA	2				
CPACT	3	CNPGL	2				
CNPS	2	CTAA	2				
CNPT	2	CNPMF	2				
CNPUV	2	CPPSUL	1				
CNPF	2	CPPSE	1				
CNPMF	2	SIÈGE	1				
CPATC	2	CNPDIA	1				
CNPAT	1	CNPA	1				
CPPSUL	1						
Total	186	Total	170				

Les centres de l'Embrapa avec le plus grand penchant international sont, en ordre d'apparition : L'Embrapa Cerrados (CPAC, DF), l'Embrapa Ressources Génétiques et Biotechnologie (Cenargen, DF), l'Embrapa Riz et Haricot (CNPAP, GO), le Siège de l'Embrapa (DF), l'Embrapa Instrumentation Agricole (CNPDA, SP), l'Embrapa Amazonie Orientale (CPATU, PA) et l'Embrapa Bovins de Viande (CNPVC, MS).

Les groupes EUA/Canada et Europe

La même analyse à été faite pour les groupes EUA/Canada et Europe. Les résultats sont présentés aux tableaux 40 et 41.

Tableau 40 : Base WOS ; Les 13 institutions de l'EUA/Canada avec plus de participations en articles par année

EUA/Canada	2002	2001	2000	1999	1998	Total
USDA-ARS	10	8	6	10	3	37
UFLORIDA	5	2	17	4	7	35
UCAL	4	8	11	6	2	31
UWISCMAD	3	1	4	9	2	19
UCORNEL		3	5	6	2	16
NCSU	1			1	11	13
MSU	3		6	1	1	11
AAFCAN	3	1	2	2		8
TEXASAM			4		4	8
CSU	2	1	3	1		7
UARIZ			1	3	1	5
UTAHSU				3	2	5
Total 13 premières	31	24	59	46	35	195
Total de participations	76	56	65	40	39	276

L'**Agricultural Research Service**, USDA-ARS (37), l'Université de la Florida, UFLORIDA (35), l'Université de la Californie, UCAL (31), l'Université de Wisconsin-Madison, UWISCMAD (19) et l'Université de Cornell, UCORNEL (16) sont évidemment dans ce groupe. La croissance de l'USDA-ARS et/ou le déclin des participations en articles d'institutions comme l'UCAL et l'UFLORIDA à partir de 2000 pourrait être attribué à un "effet LABEX EUA". Le premier Laboratoire Virtuel de l'Embrapa à l'Extérieur (LABEX) à été créé le 28 avril, 1998 aux États Unis, en coopération avec l'USDA-ARS. Les travaux conjoints de recherche développés depuis cette date ont contribué à élever le score de l'USDA-ARS en relation aux autres institutions américano-canadiennes.

Un effet similaire serait rencontré avec des institutions françaises (CIRAD et INRA) après l'établissement du LABEX Europe, le 23 janvier, 2001, dans l'Agropolis, à Montpellier. Cette évolution pourrait être accompagnée par des études futures. Ces détails sont montrés dans le tableau 41.

Tableau 41 : Base WOS ; Les 13 premiers partenaires de l'Embrapa en Europe par année

Europe	2002	2001	2000	1999	1998	Total
UWAGENINGEN	8	2	5	7	1	23
UBAYREU	2	2	11	5	2	22
CIRAD/IRD/ORSTOM	6	3	4	4	3	20
ULONDON	3	2	4		2	11
INRA	4	2	1	2	2	11
UHAMBURG	6		2	1	1	10
UCATLOUVAIN	4		5			9
UGOETTINGEN	4		1	2	1	8
UVALENCIA	2	1	1	2	2	8
UREADING	1	4			1	6
GSF		2	1	1	1	5
ULEEDS	2		2		1	5
UATHENS	2	3				5
Total 13 premières	44	21	37	24	17	143
Total de participations	76	49	59	44	21	249

L'Université de Wageningen, UWAGENINGEN (23), l'Université de Bayreuth, UBAYREU (22), le Centre de Coopération Internationale en Recherche Agronomique pour le Développement, CIRAD/IRD/ORSTOM (20), l'Institut Nationale de Recherche Agronomique, INRA (11), l'Université de Londres, ULONDON (11) et l'Université de Hambourg, UHAMBURG (10) se détachent en Europe. À noter que la somme des productions scientifiques des institutions françaises représenterait la première position.

Les groupes Afrique/Amérique Latine/Asie et Système ONU

La même analyse a été faite pour les groupes Afrique/Amérique Latine/Asie et Système ONU. Les résultats sont présentés dans les tableaux 42 et 43.

Tableau 42 : Base WOS ; Les dix premiers partenaires de l'Embrapa en Afrique/Amérique Latine/Asie par année

Afrique/Amérique Latine/Asie	2002	2001	2000	1999	1998	Total
CSIRO (Australie)	2		2	1	1	6
UTOKYO (Japon)		1			4	5
ULINCLN-NZ (Nouvelle Zélande)	1		2		1	4
UNAM (Mexique)		1	1	1	1	4
AHMADBUN (Nigeria)	1	1		2		4
UHOKKAIDO (Japon)	2			1		3
UQUEENSLAND (Australie)	1		1			2
UNASSUMP-PAR (Paraguay)	1		1			2
UNACNE-ARG (Argentine)	1				1	2
UNACSUR-ARG (Argentine)				1	1	2
Autres AFR/AMLAT/ASIE	9	4	5	2	4	24
Total	18	7	12	8	13	58

Le critère de privilégier les institutions avec deux ou plus participations en articles a réduit la présence ostensive des institutions de l'Amérique Latine. Le CSIRO de l'Australie et l'Université de Tokyo du Japon sont les deux premières institutions de ce groupe suivis par l'Université de Lincoln, de la Nouvelle Zélande, l'Université Nationale Autonome du Mexique et l'Université Ahmadu Bello, du Nigeria. Quant au Système ONU, le principal partenaire est le Centre Internationale de l'Agriculture Tropicale (CIAT) basé au Costa Rica (Amérique Centrale).

Tableau 43 : Base WOS ; Les partenaires de l'Embrapa en le Système ONU par année

Système ONU	2002	2001	2000	1999	1998	Total
CIAT	1	3	6	5	3	18
CIFOR	2			1		3
ICRISAT	3					3
Total	6	3	6	6	3	24

L'étude des revues scientifiques

Nous avons décidé d'étudier les participations dans la production d'articles par les revues scientifiques. Le tableau 44 extrait les revues qui ont divulgué jusqu'à huit articles des chercheurs de l'Entreprise.

Tableau 44 : Base WOS ; Participation en articles par revue scientifique

Revues scientifiques	2002	2001	2000	1999	1998	Total	%
PESQUISA AGROPECUARIA BRASILEIRA	85	52	122	111	78	448	27,5
REVISTA BRASILEIRA DE ZOOTECNIA-BRAZILIAN JOURNAL OF ANIMAL SCIENCE	24	27	24	53	37	165	10,1
ARQUIVO BRASILEIRO DE MEDICINA VETERINARIA E ZOOTECNIA	12	12	14	6	11	55	3,3
GENETICS AND MOLECULAR BIOLOGY	12	2	20	9	10	53	3,2
COMMUNICATIONS IN SOIL SCIENCE AND PLANT ANALYSIS	10	9	5	4	5	33	2,2
PESQUISA VETERINARIA BRASILEIRA	3	6	8	7	6	30	1,8
EUPHYTICA	5	8	2	5	2	22	1,3
BRAZILIAN ARCHIVES OF BIOLOGY AND TECHNOLOGY	6	4	4	5	3	22	1,3
JOURNAL OF PLANT NUTRITION	3	1	1	5	8	18	1,1
PLANT AND SOIL	4	3	5	1	5	18	1,1
BIOLOGY AND FERTILITY OF SOILS	3	2	6		3	14	0,8
BRAZILIAN JOURNAL OF MICROBIOLOGY	5	5	4			14	0,8
FOREST ECOLOGY AND MANAGEMENT	4	4	4	1		13	0,7
THEORETICAL AND APPLIED GENETICS	2	3	3	3	1	12	0,7
THERIOGENOLOGY	6		1	1	3	11	0,6
PROTEIN AND PEPTIDE LETTERS	4	2	3	1		10	0,6
MEMORIAS DO INSTITUTO OSWALDO CRUZ	1	2	1	2	3	9	0,5
GEODERMA	2	1	2	3	1	9	0,5
VETERINARY PARASITOLOGY		1	1	3	4	9	0,5
JOURNAL OF APPLIED ENTOMOLOGY-ZEITSCHRIFT FUR ANGEWANDTE ENTOMOLOGIE	1		3	2	2	8	0,5
ENVIRONMENTAL ENTOMOLOGY	1	3	4			8	0,5
REVISTA DE MICROBIOLOGIA				6	2	8	0,5
Total 22 premières revues	193	147	237	228	184	989	60,1
Total de participations	354	273	374	355	272	1628	100

La revue scientifique préférée des chercheurs du groupe Embrapa est "PESQUISA AGROPECUARIA BRASILEIRA" qui concentre 27,5% du total d'articles. Autre publication importante est la "REVISTA BRASILEIRA DE ZOOTECNIA-BRAZILIAN JOURNAL OF ANIMAL SCIENCE" qui publie 10,1% des articles. Ces deux revues et la prochaine de la liste, "ARQUIVO BRASILEIRO DE MEDICINA VETERINARIA E ZOOTECNIA", avec 3,3% des articles, réunissent plus du tiers de la production scientifique de plusieurs centres de recherche de l'Embrapa, classés par le total de participations en articles publiés. Pour 19 centres, indiqués en gris au tableau 45, la proportion de participations dans les trois revues dépasse 50%. Pour neuf d'autres ces participations sont plus de 33% comme le montre le tableau 45.

Tableau 45 : Base WOS ; Les trois revues les plus préférées par les centres Embrapa

Revues scientifiques \ Centres	PESQUISA AGROPECUARIA BRASILEIRA	REVISTA BRASILEIRA DE ZOOTECNIA- BRAZILIAN JOURNAL OF ANIMAL SCIENCE	ARQUIVO BRASILEIRO DE MEDICINA VETERINARIA E ZOOTECNIA	Total des trois revues	Total de participations du centre	% participations du centre
1 CPAF-RR	2			2	2	100
2 CAFE		1		1	1	100
3 CPAO	12	1		13	14	92,8
4 CPAF-RO	5			5	6	83,3
5 CNPGL	15	47	26	88	106	83
6 CPATSA	25	19		44	54	81,5
7 CPAMN	4			4	5	80
8 CPATC	19			19	24	79,2
9 CNPT	27	2		29	40	72,5
10 CNPSA	13	26	8	47	66	71,2
11 CNPC	5	9	5	19	28	67,8
12 CNPA	13			13	21	61,9
13 CNPTIA	3			3	5	60
14 CPAF-AP	3			3	5	60
15 CPAF-AC	3	1		4	7	57,1
16 CNPMF	13			13	23	56,5
17 CNPGC	6	27		33	61	54,1
18 CNPAF	41	1		42	81	51,8
19 CNPUV	7			7	14	50
20 SIÈGE	5	4		9	19	47,4
21 CNPAT	15			15	33	45,4
22 CPPSE	6	26	1	33	74	44,6
23 CNPS	16			16	36	44,4
24 CNPMS	22	2	7	31	70	44,3
25 CPAC	23	5	3	31	71	43,7
26 CPACT	5	2		7	17	41,2
27 CPATU	17			17	44	38,6
28 CPAP	13	4	1	18	47	38,3
29 CPAA	11			11	35	31,4
30 CNPSO	24			24	81	29,6
31 CNPH	10			10	38	26,3
32 CNPMA	18			18	72	25
33 CPPSUL	1	1		2	8	25
34 CTAA	6			6	29	20,7
35 CNPF	2			2	16	12,5
36 CNPAB	8			8	65	12,3
37 CENARGEN	13	4	1	18	224	8
38 CNPDIA	1			1	75	1,3
39 CNPM				0	1	0
Total de participations	432	182	52	666	1618	41,16

Cette information confirme le fait que la base Pascal, comme elle n'indexe aucune de ces revues, laisse de côté une partie significative de la production scientifique brésilienne en agriculture. A remarquer que l'Embrapa Ressources Génétiques et Biotechnologie (Cenargen) et l'Embrapa Instrumentation Agricole (CNPDIA), marqués en gras, publient moins de 10% de leurs articles dans ces trois revues.

L'étude des langues

Dans la base WOS, où la représentativité des périodiques scientifiques brésiliens est plus grande que dans la base Pascal, la tendance déjà perçue de certains centres de recherche de publier une partie significative de leur production scientifique en portugais s'est accentuée et caractérisée. Ceci est le cas, d'entre autres, de l'Embrapa Bovins de Lait (CNPGL), l'Embrapa Porcs et Volailles (CNPSA), l'Embrapa Semi-Aride (CPATSA), l'Embrapa Bovins de Viande (CNPGC), l'Embrapa Blé (CNPT), l'Embrapa Caprins (CNPC) et l'Embrapa Plateaux Côtiers (CPATC), d'entre autres. Ils sont marqués en gris dans le tableau 46. A remarquer une particularité singulière. L'Embrapa Bovins de Viande (CNPGC) qui est signalée en gras, qui est un des sept centres plus internationaux de l'Embrapa publie la majorité de ses travaux en portugais. Deux explications contrastantes sont en vue : il n'est pas impossible d'avoir des relations internationales étendues en parlant portugais ou encore les relations internationales de l'Embrapa sont encore peu développées. Les deux nous semblent valables.

Tableau 46 : Base WOS ; Groupe Embrapa, participation en articles par centre et par langue

Centre	Anglais	Portugais	Espagnol	Français	Allemand	Total
1 CNPGL	20	86	1			107
2 CNPSA	25	43		1		69
3 CPATSA	12	43				55
4 CNPGC	25	40	1			66
5 CPPSE	37	36	1			74
6 CNPAF	47	33		1		81
7 CPAC	43	29		1		73
8 CNPT	13	27				40
9 CNPMS	40	27	3			70
10 CNPC	8	20				28
11 CPATC	5	19				24
12 CNPMA	56	17				73
13 CPAP	37	15				52
14 CNPS	19	14	2	1		36
15 CNPAT	19	14				33
16 CPAO		14				14
17 CENARGEN	213	13	1			227
18 CNPSO	68	13				81
19 CNPMF	10	12	1			23
20 CNPAB	54	11				65
21 CNPA	10	11				21
22 CPATU	32	10	2			44
23 CPAA	24	9	1		1	35
24 SIÈGE	11	8				19
25 CNPH	29	8	1			38
26 CNPUV	6	7		1		14
27 CPACT	9	7			1	17
28 CTAA	22	6	1			29
29 CPAF-RO	1	5				6
30 CPAF-AC	3	4				7
31 CNPTIA	2	3				5
32 CPAF-AP	2	3				5
33 CNPF	14	2				16
34 CPPSUL	6	2				8
35 CPAF-RR		2				2
36 CNPDIA	73	2				75
37 CPAMN	2	2	1			5
38 CAFE		1				1
39 CNPM	1					1
Total	998	618	16	5	2	1639

Conclusion

Le travail d'audit de la Collection documentaire et de création d'indicateurs de la production scientifique de l'Embrapa est une démonstration des traitements et combinaisons possibles de données dans une base de données bibliographique, une contribution pour l'amélioration de la qualité de toute la Collection documentaire et pour son intégration dans le système d'informations gestion de l'Embrapa, par la création d'indicateurs de production scientifique directement à partir de bases primaires.

Le travail d'analyse des bases de l'Embrapa, WOS et Pascal a mis en évidence l'intérêt du champ d'affiliation des auteurs. Ceci est l'un des différentiels de ces deux bases internationales. Les informations d'affiliation sont pour les analystes d'informations textuelles comme de l'or pour des miniers. Elles permettent de dessiner des réseaux d'auteurs et de collaborations et d'analyser leur évolution dans le temps. Les analyses technologiques et de brevets sont les principales consommatrices de ce genre d'information.

En plus d'être aussi des critères de qualité de bases de données SCOUG (Medawar, 1995 ; Basch 1995, Faria, 2001), il serait opportun que lors du processus en cours à l'Embrapa de création de champs spécifiques pour le type de document (pour la génération automatique d'indicateurs de production scientifique) soient créés aussi : (1) Champs spécifiques pour les informations sur les revues scientifiques, avec la classification de leur champ scientifique et (2) Champs d'affiliation pour tous les auteurs, en suivant les recommandations de Faria (2001, p.98) en gras :

- *informations de l'article ; Titre, résumé, mots-clés libres attribués à l'article, **mots-clés contrôlés attribués à l'article, langue, type de document** (article scientifique, livre, communication en congrès, ...), année de publication;*
- *informations de la revue ; **Nom du périodique dans lequel l'article a été publié, classification du champ de la connaissance du périodique;***
- *informations des auteurs ; **Auteurs, entités d'affiliation, départements de travail, ville et pays.***

Finalement il serait fortement souhaitable que les institutions scientifiques, d'une manière générale, normalisent les données d'affiliation des auteurs avant la soumission des articles à la publication pour éviter des erreurs de comptage quand de la construction des divers indicateurs de production scientifique.

Analyse stratégique des compétences de l'Embrapa

Le développement des bases de compétences

Dans le travail intitulé "De la création des bases de données au développement de systèmes d'intelligence pour l'entreprise", Penteado et al. (2003a) démontrent que les bases de données opérationnelles existantes dans les organisations peuvent devenir des mines d'informations pour leurs directeurs, pourvue qu'elles soient structurées de manière à permettre des analyses plus élaborées et affirment "qu'en effet, stocker des données pour une utilisation faisant simplement appel à 10 ou 20% des utilisations possibles conduit à une perte de temps et aussi à un affaiblissement stratégique" (p.6). Et continuent (Penteado et al., 2003a, p.2) :

Actuellement, l'accroissement constant des informations à la fois au plan international, mais aussi au sein même de l'entreprise pose le problème de la manière dont ces informations vont être construites, associées, traitées et utilisées à des fins opérationnelles de recherche, de valorisation, de conquête commerciale. Dans le travail qui est présenté ici, nous allons à partir de plusieurs exemples, mettre en évidence les utilisations de l'information dans le cas du système de renseignement OUVERT de l'entreprise. (...)

Pour Penteado et al. (2003a, p.2), le processus de création de l'information stratégique consiste en des étapes de

Collecte de l'Information, Indexation et Gestion - Gestion des Compétences - Détermination des Spécialistes - Organisation du Travail des Spécialistes. Nous allons dans ce travail focaliser notre attention sur ces différents points, entre autres à partir de la base des compétences des chercheurs de l'Embrapa, en mettant en évidence chaque fois que cela sera possible les traitements nécessaires au niveau de la base elle-même, puis les résultats pouvant être obtenus à partir d'un traitement de la base par des logiciels bibliométriques spécialisés.

Nous traiterons, en particulier, de la gestion des compétences (Diplômes et âges), de la capacité à transférer des technologies, puis des spécialisations en recherche.

L'ACCES AUX INFORMATIONS, LA BASE DE DONNEES, SA STRUCTURE ET TRAITEMENTS

Si, au niveau de l'Intelligence Compétitive, il faut de préférence collecter des informations formelles et informelles, nous mettrons ici l'accent sur les informations formelles issues de la base des compétences des chercheurs de l'Embrapa.

La base de données, telle qu'elle nous a été fournie est structurée de la manière suivante.

Tableau 47 : Structure originale de la base Compétences (Penteado et al., 2003a, p.2)

NOME : JOSE IVO BALDANI
DATN : 12\2\1953
FORM : AGRONOMIA, 1976; MESTRADO: CIENCIA DO SOLO, UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO, 1984; DOUTORADO: CIENCIA DO SOLO, UNIVERSIDADE DO TEXAS A&M - ESTADOS UNIDOS, 1990.
APSQ : BIOLOGIA MOLECULAR; BACTERIAS FIXADORAS DE NITROGENIO; CONTROLE BIOLOGICO; GRAMINEAS
PROD : BIOINSETICIDAS, BIOFERTILIZANTES, BIOTECNOLOGIA
UNID : EMBRAPA AGROBIOLOGIA

Les champs documentaires sont cités au tableau suivant.

Tableau 48 : Les champs documentaires (Penteado et al., 2003a, p.2)

NOME = Nom (Origine = Guide de Sources Embrapa)
DatN = Date de naissance (Origine = Base RH Embrapa)
FORM = Formation (Origine = Guide de Sources Embrapa)
APSQ = Thème de recherche (Origine = Guide de Sources Embrapa)
PROD = Produits (Origine = Guide de Sources Embrapa)
UNID = Centre de Recherche de l'Embrapa (Origine = Guide de Sources Embrapa)

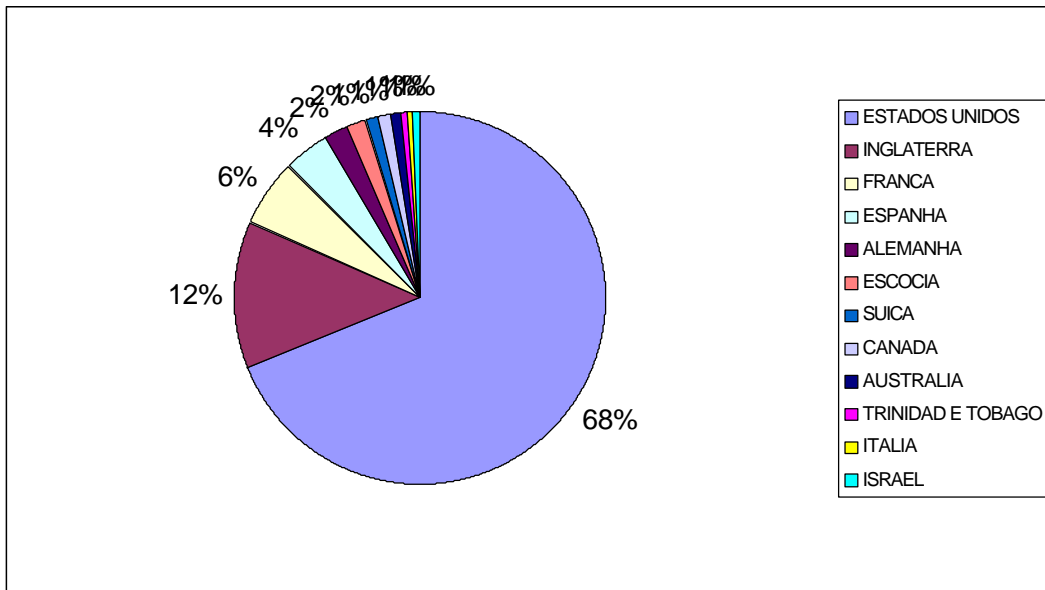
Nous avons constaté l'absence d'un certain nombre de champs documentaires pouvant être mis en place à partir de la base initiale et permettant ainsi un traitement plus rapide et plus en adéquation avec l'Intelligence Compétitive et/ou la gestion des ressources humaines. Par exemple le champ COOR ou Cohorte a été créé. Il permet, à partir de la date de naissance de diviser le personnel en 5 Cohortes : N°1 moins de 25 ans - N°2 entre 25 et 34 ans - N°3 entre 35 et 44 ans - N°4 entre 45 et 54 ans - N°5 plus de 55 ans. En outre, nous avons éclaté le champ formation dans des différents champs : Type de Diplôme, année d'obtention, lieu d'obtention. Nous avons constaté dans ce cas de nombreuses déviations quant à l'écriture ce qui ne permet pas de travailler statistiquement et ce qui montre que l'utilisation de la base de données était faite sans valeur ajoutée. Par exemple, on peut à partir du nom d'une personne connaître sa formation, mais on ne peut pas connaître toutes les personnes ayant une formation du même type, à cause des non-homogénéité dans la saisie des données. Les données de départ ont donc du être reformatées. Un exemple de référence reformatée est présenté au tableau 49.

Tableau 49 : Structure modifiée de la base Compétences (Penteado et al., 2003a, p.3)

NOME : JOSE IVO BALDANI
DATN : 12\2\1953
COOR : 4
FORM : AGRONOMIA formation initiale
QDFORM:1976 date de la formation initiale
MESTRADO: CIENCIA DO SOLO type de Diplôme et spécialité
QDMEST:1984 date d'obtention du diplôme
UNIVMEST: UNIVERSIDADE FEDERAL RURAL DO RIO DE JANEIRO Université d'obtention
DOUTORADO: CIENCIA DO SOLO type de Diplôme et spécialité
DOUEST : 1990 dates d'obtention du diplôme
UNIVDOU : UNIVERSIDADE DO TEXAS A&M Université d'obtention
PAISDOU : ESTADOS UNIDOS Pays d'obtention
APSQ : BIOLOGIA MOLECULAR; BACTERIAS FIXADORAS DE NITROGENIO; CONTROLE BIOLOGICO; GRAMINEAS
PROD : BIOINSETICIDAS, BIOFERTILIZANTES, BIOTECNOLOGIA
UNID : EMBRAPA AGROBIOLOGIA

Une fois ce retraitement effectué, on peut réaliser différents types d'analyses, à partir des différents champs. On peut par exemple s'intéresser au doctorat : Répartition par pays, graphes relatifs aux dates d'obtention, corrélation avec la pyramide des âges de l'Embrapa, différences de répartition des tranches d'âge entre unités. Nous allons à titre d'exemple, présenter dans la figure 12, la répartition des doctorats par pays, puis dans la figure 13, la pyramide des âges de l'Embrapa et dans la figure 14, la corrélation entre pyramide des âges et l'année d'obtention des doctorats, au Brésil et à l'étranger.

Le nombre de docteurs étrangers est de 190, le nombre de docteurs brésiliens est de 286.



Note. [Etats-Unis, Angleterre, France, Espagne, Allemagne, Écosse, Suisse, Canada, Australie, Trinidad et Tobago, Italie, Israël]

Figure 12 : Répartition des doctorats passés à l'étranger (Penteado et al., 2003a, p.3)

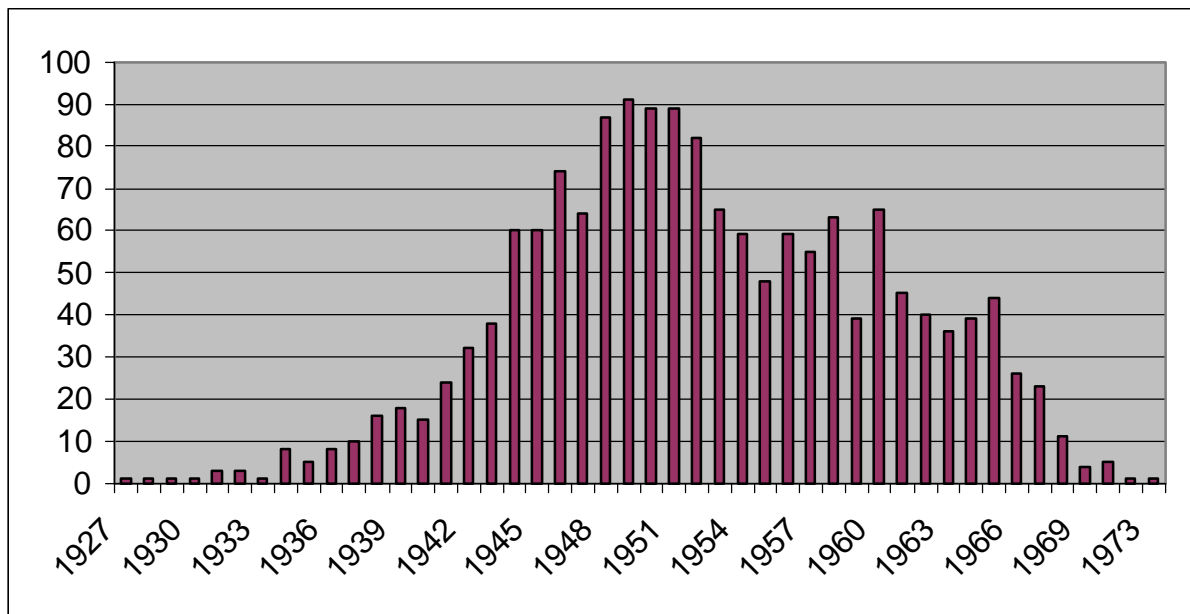


Figure 13 : Pyramide des âges de l'Embrapa (Penteado et al., 2003a, p.3)

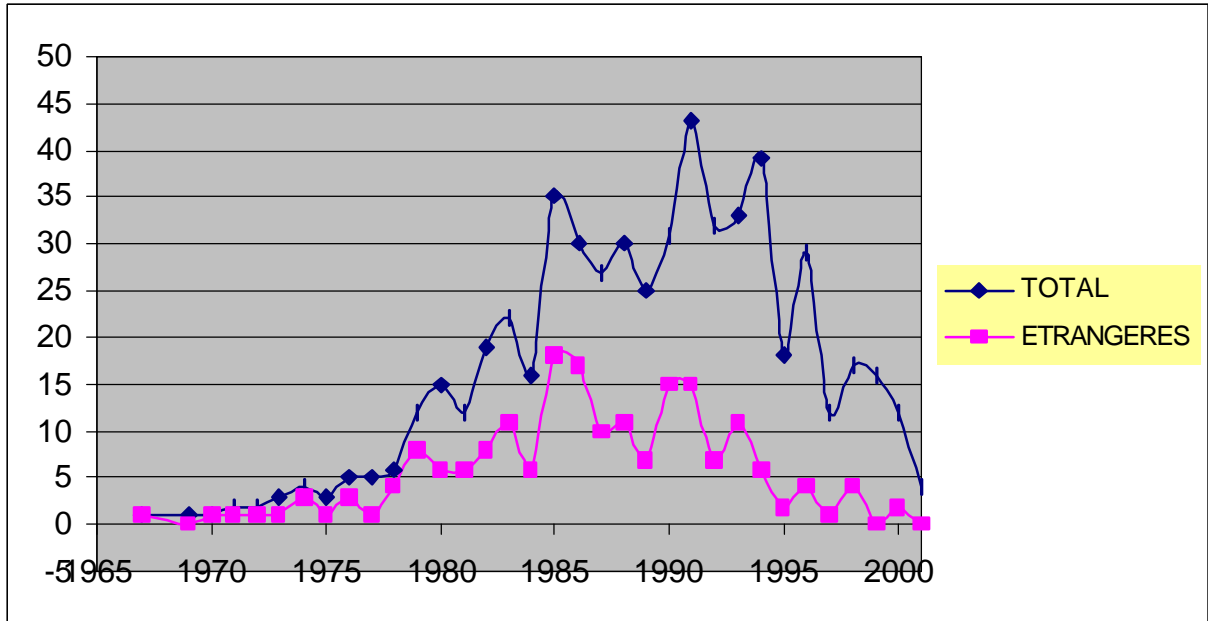


Figure 14 : Personnels de l'Embrapa, répartition des thèses de Doctorat par an (Penteado et al., 2003a, p.3)

On constate que ce graphe est assez proche de celui de la pyramide des âges et que le ratio docteurs brésiliens / docteurs étrangers est relativement homogène.

D'autres façons de graphes peuvent être présentées si on s'intéresse par exemple à l'université ou à la formation initiale des collaborateurs. Ces analyses peuvent par exemple servir de base à la création d'un groupe d'experts dans lesquels on évitera de créer des groupes à forte consanguinité ayant une "pensée unique", du fait de l'origine des formations.

Boutin (2001) a utilisé cette même base pour étudier le niveau de consanguinité des chercheurs de l'Embrapa Cerrados. Voyez dans la Figure 17.

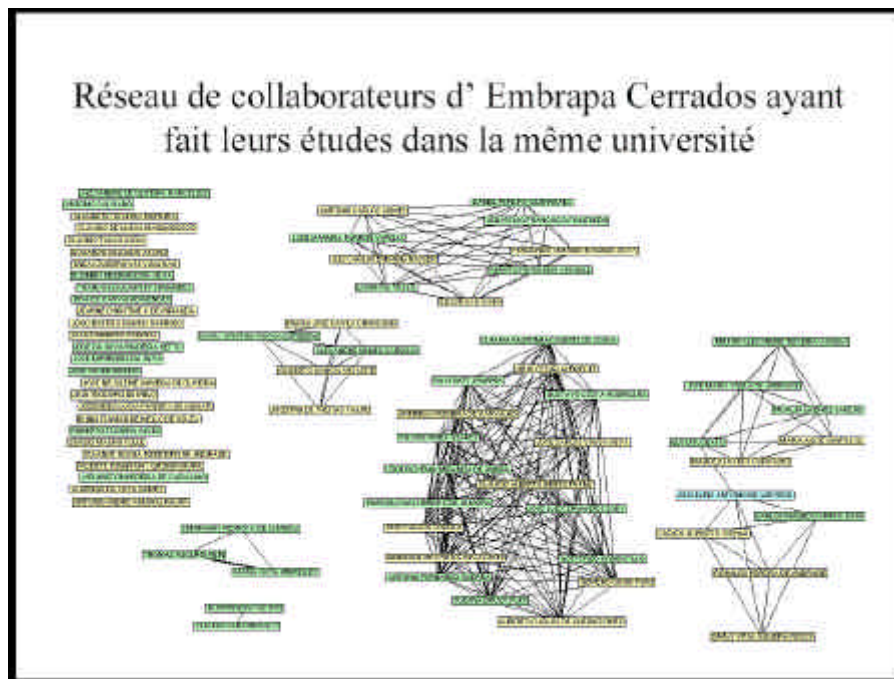


Figure 15: Étude de la consanguinité des chercheurs de l'Embrapa Cerrados (Boutin, 2001, p.7)

Et continuent Penteadó et al. (2003a):

Il est aussi possible, puisque nous avons créé le champ cohorte, de mettre en évidence les différences de profil des âges des personnels entre les différentes unités de l'Embrapa. Ceci est réalisé rapidement en croisant les champs cohorte et unité d'appartenance. On arrive alors, après traitement, aux résultats suivants.

COHORTE 2

Dans la cohorte entre 25 et 34 ans, il n'y a pas de différences significatives entre les unités de l'Embrapa sauf dans quatre d'entre elles pour lesquelles le nombre de personnes de la cohorte 2 est significativement supérieur à la moyenne.

EMBRAPA UVA E VINHO

EMBRAPA INFORMATICA AGROPECUARIA

EMBRAPA RORAIMA

EMBRAPA ACRE

Cela peut s'expliquer par :

Des axes stratégiques pour l'entreprise : Il serait intéressant de regarder de plus près la production de ces personnes.

Des départs de collaborateurs qu'il a fallu remplacer.

Des raisons politiques.

Des problèmes de pyramide des âges dans l'entreprise.

COHORTE 3

EMBRAPA AMAZONIA ORIENTAL

EMBRAPA SEDE

Dans ces deux unités il y a moins de personnes dans la cohorte 3 que la moyenne.

EMBRAPA AMAZONIA OCIDENTAL

EMBRAPA SOLOS

EMBRAPA SUINOS E AVES

EMBRAPA ACRE

EMBRAPA PANTANAL

EMBRAPA INSTRUMENTACAO AGROPECUARIA

EMBRAPA AMAPA

Dans ces sept unités, il y a plus de personnes dans la cohorte 3 que la moyenne.

COHORTE 4

EMPARN

Dans cette unité, il y a plus de personnes de 40 ans que la moyenne.

COHORTE 5

EMBRAPA AMAZONIA ORIENTAL

EMBRAPA CLIMA TEMPERADO

EMBRAPA SEDE

Dans ces trois unités, il y a plus de personnes dans la cohorte 5 que la moyenne. Donc embauche probable à moyen terme. À souligner qu'au siège des entreprises, d'une manière générale (EMBRAPA SEDE) il est naturel d'y trouver des personnels plus âgés.

EMBRAPA MEIO AMBIENTE

EMBRAPA SUINOS E AVES

Dans ces deux unités, il y a moins de personnes dans la cohorte 5 que la moyenne.

LE DEVELOPPEMENT PRODUIT (TRANSFERT DE TECHNOLOGIES, ECONOMIE)

L'objectif de toute institution est bien de favoriser son interaction avec le monde économique afin de valoriser ses produits, de mettre en œuvre des politiques de transfert de technologies et de créer les meilleures conditions possibles au plan des négociations commerciales. L'analyse que nous conduirons utilise le champ produit (PROD). Ce champ nous permet de savoir quels sont les produits ou activités à caractère transférable qui ont fait l'objet d'une saisie pour les individus présents dans la base. Ceci signifie que ces personnes ont une compétence dans les domaines économiques indiqués (à ne pas confondre avec les données concernant les thèmes de recherche qui est une approche plus fondamentale et non directement liée à court terme à la valorisation).

Notre objectif est le suivant : Mettre en évidence la capacité générale de transfert économique de l'Embrapa. Cela va se faire par la construction d'une matrice carrée (réalisation de toutes les paires de termes contenues dans le champ PROD. À noter entre autres qu'il n'y a que peu de personnes (moins de 10%) qui n'ont qu'une seule spécialité dans ce cadre. Une telle matrice carrée est réalisée rapidement à partir de logiciels bibliométriques spécialisés (nous donnerons en annexe quelques détails sur ces derniers). Une fois la matrice réalisée, elle va être transformée en une représentation infographique qui va permettre, visuellement, d'appréhender les capacités économiques de transfert de l'Embrapa (basée sur ses propres spécialistes).

Cette représentation des réseaux des associations de domaine de recherche des cohortes 2 et 3 est présentée dans la figure 16.

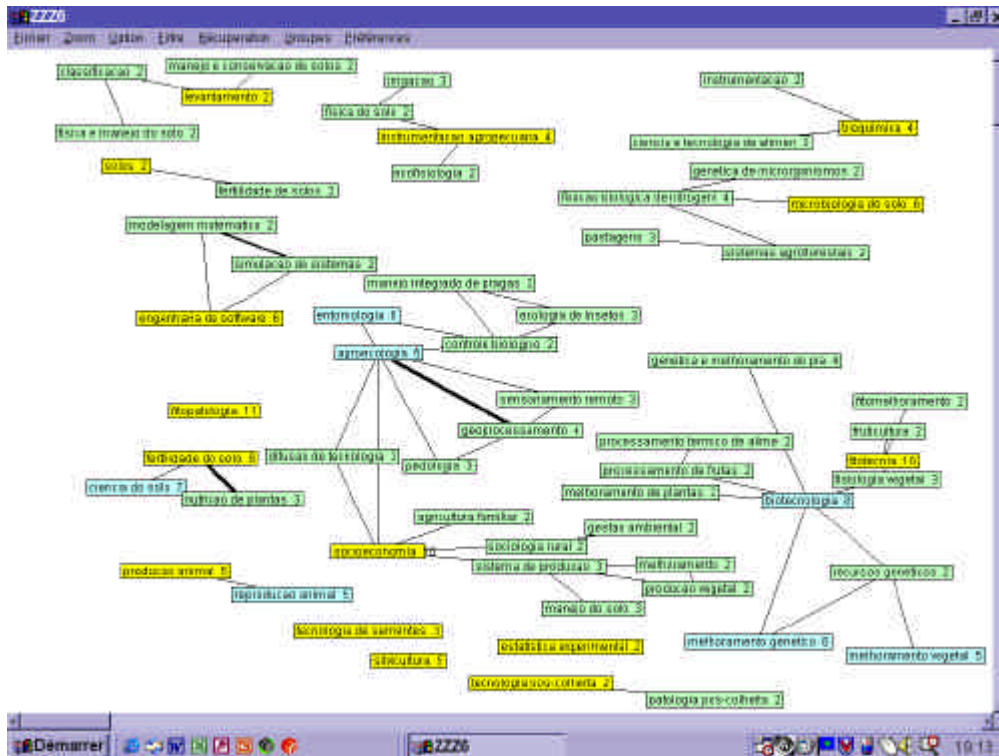


Figure 16 : Réseau des compétences économiques par produits de l'Embrapa (Penteado et al., 2003a, p.4)

Le réseau est présenté ici dans sa globalité. Que veut dire le fait que deux domaines soient liés ? Ceci veut dire qu'il existe au sein de l'Embrapa, au moins un chercheur qui maîtrise ces deux compétences. Sur le plan de la stratégie, cette cartographie a l'avantage de préciser l'ensemble des compétences économiques et en même temps les liens entre elles. On peut donc, lorsqu'on va aborder un problème, ou lorsqu'on veut réaliser un transfert, ou répondre à une demande, exprimer cela en fonction des mêmes termes produits que ceux utilisés dans la base de données et dans la cartographie et positionner la demande ou le transfert sur cette carte. On verra ainsi, dans quelle mesure on peut répondre à la question et avec quels efforts.

Pour donner plus de détails, nous avons zoomé sur une partie présentée Figure 17.

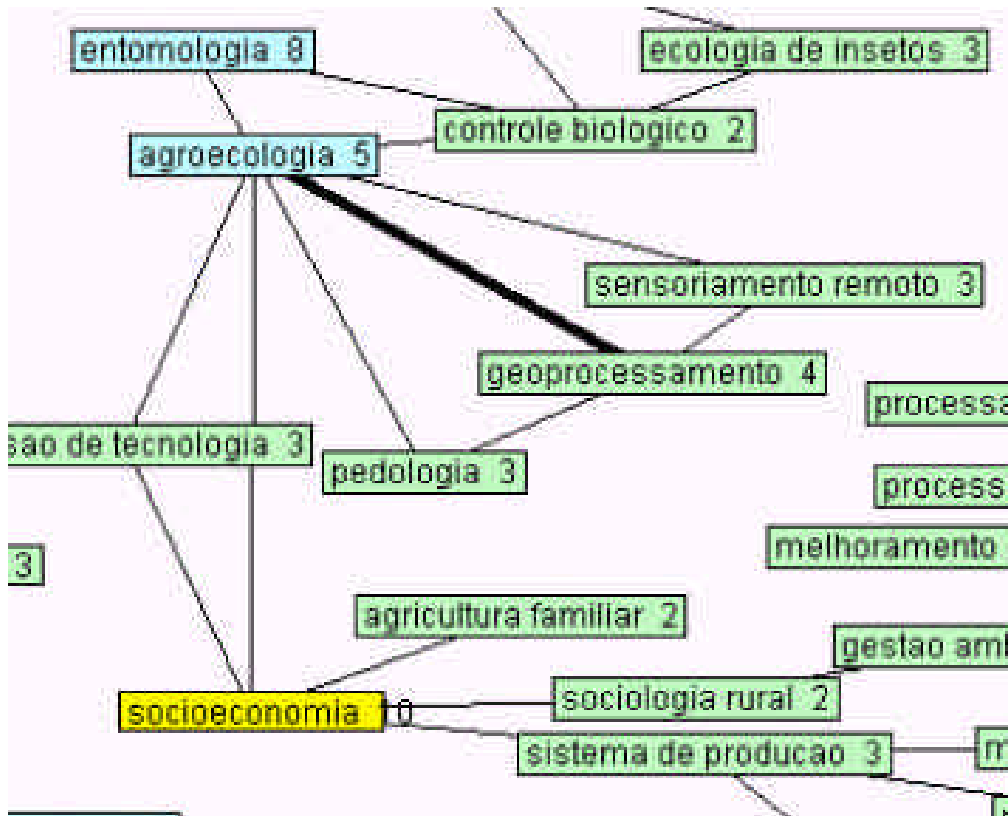


Figure 17 : Détail du réseau précédent (Penteado et al., 2003a, p.5)

On constate ici les liens entre les différentes capacités de transfert de l'Embrapa. Certains liens sont forts, d'autres sont plus faibles. Certains thèmes récents comme SOCIOECONOMIA, sont souvent des nœuds du réseau. D'autres plus anciens comme AGROECOLOGIA constituent de points centraux à partir desquels des processus de transferts pourront être établis.

Reste maintenant, à descendre au niveau des spécialistes. En effet, la cartographie précédente ne fait que mettre en évidence la capacité de réponse de l'Embrapa, mais ne détermine pas avec quels spécialistes. Pour ce faire, nous allons travailler sur les multi-spécialités de ces derniers. Il faut alors réaliser un traitement bibliométrique par paires en associant le champ NOME (nom des personnes) à celui des compétences. On classera ensuite le tout par ordre alphabétique et on aura alors, un accès rapide aux multi-compétences. Ceci est indiqué dans le tableau suivant.

Tableau 50 : Exemple de tableau multi-compétences - Embrapa (Penteado et al., 2003a, p.5)

NOM	COMPÉTENCES
ADRIANO VENTURIERI	ZONEAMENTO
ADRIANO VENTURIERI	SENSORIAMENTO REMOTO
ADRIANO VENTURIERI	LEVANTAMENTOS FLORESTAIS
ADRIANO VENTURIERI	ANALISE AMBIENTAL
ALBERTO WILLIAM VIANA DE CASTRO	TRANSFERENCIA DE TECNOLOGIA
ALBERTO WILLIAM VIANA DE CASTRO	POLITICA DE CIENCIA E TECNOLOGIA
ALBERTO WILLIAM VIANA DE CASTRO	PLANEJAMENTO ESTRATEGICO
ALBERTO WILLIAM VIANA DE CASTRO	GESTAO DA INOVACAO
ALFREDO KINGO OYAMA HOMMA	SISTEMAS AGROFLORESTAIS
ALFREDO KINGO OYAMA HOMMA	SILVICULTURA
ALFREDO KINGO OYAMA HOMMA	PIMENTA-DO-REINO
ALFREDO KINGO OYAMA HOMMA	ECOLOGIA
ALFREDO KINGO OYAMA HOMMA	AGROFLORESTA
ALTEVIR DE MATOS LOPES	ARROZ RECURSOS GENETICOS
ALTEVIR DE MATOS LOPES	ARROZ MELHORAMENTO
ANGELA MARIA LEITE NUNES	PIMENTA-DO-REINO
ANGELA MARIA LEITE NUNES	FRUTICULTURA PRAGAS
ANGELA MARIA LEITE NUNES	CUPUACU
ANGELA MARIA LEITE NUNES	CONTROLE DE PRAGAS
ANTONIO AGOSTINHO MULLER	LEGUMINOSAS
ANTONIO AGOSTINHO MULLER	DENDE
ANTONIO AGOSTINHO MULLER	CASTANHA-DO-BRASIL
ANTONIO DE BRITO SILVA	MANEJO INTEGRADO DE PRAGAS
ANTONIO DE BRITO SILVA	CONTROLE DE PRAGAS

Ceci est réalisé pour l'ensemble des personnels de l'Embrapa présents dans la base. Pour utiliser de tels résultats sur le plan stratégique, cela va se faire soit au plan des groupes d'experts (dans le choix de ces derniers), soit dans le choix des personnes pouvant participer à une négociation. Supposons que nous voulions transférer des compétences en systèmes agroforestiers à un Pays qui a déjà des compétences en sylviculture ou des deux. La personne qui va négocier doit-elle : Être uniquement spécialiste en systèmes agroforestiers, ou de sylviculture, ou des deux. Dans le premier cas on fera uniquement appel à un spécialiste du domaine, mais dans l'autre cas à une double compétence. C'est en ce sens que

l'identification rapide des multi-compétences devient un outil stratégique de premier plan.

Un autre exemple peut être présenté. Considérons par exemple, une entreprise réalisant divers arômes pour des sorbets, des crèmes des gâteaux. Le problème est de réaliser au sein de l'entreprise des innovations dans ce domaine. Comment peut-on procéder ?

L'approche classique est de placer dans l'entreprise, une personne qui sera spécialisée dans la gestion de l'innovation et le transfert de technologie. Dans ce cas ALBERTO WILLIAM VIANA DE CASTRO sera choisi. Mais, stratégiquement l'examen par des experts de la question posée, va nous conduire à vouloir obtenir plus de cette demande, à l'orienter et compte tenu des axes choisis à faire une ou des propositions précises répondant à la question de l'innovation. Dans ce cas, l'analyse des compétences produits précédentes nous conduit à prendre en compte le CUPUACU parce que (La collecte, l'évaluation et le manège de matériel natif du cupuacu ont possibilité d'identifier des clones de cet arbre fruitier tropical avec des caractéristiques de grande utilisation pour la consommation en nature ou l'industrialisation du fruit pour son jus. (...) Ciência para a vida - Embrapa, 1998). Dans ce cas, il faudra associer à la personne précédente ANGELA MARIA LEITE NUNES.

CONCLUSION: TRANSFORMER DES BASES DE DONNÉES EN RESSOURCES STRATÉGIQUES

On voit ainsi comment l'utilisation stratégique des multi-compétences devient importante dans le cas de l'Intelligence Compétitive. Ce n'est pas notre objectif que de développer des stratégies ou des scénarios possibles, l'exemple pris ci-dessus montre simplement, comment une base de données doit être structurée ou analysée pour être utilisable au plan stratégique.

En approfondissant l'analyse des compétences

La base de compétences Embrapa utilisé par Penteado et al. (2003a) a son origine dans le Guide de Sources de l'Embrapa (2005c) qui a été combiné avec des informations comme date d'admission et date de naissance de la Base des ressources humaines de l'Embrapa (Penteado et al., 2005a, p.17). Ont été ajoutées aussi des informations d'une troisième base, celle de recherche et développement de l'Embrapa, créée par la Surintendance de Recherche et Développement (SPD). Le Guide de Sources a été ainsi décrit par Penteado et Quoniam (2001, p.6)

*à l'Embrapa, le Bureau de Communication Sociale maintient, depuis 1996, une banque de compétences des chercheurs, connu comme le "Guide des Sources"(**Guia de Fontes**) où il est possible de savoir qui fait quoi dans*

l'entreprise. Cette base a été créée initialement comme un instrument de travail pour tous les techniciens de communication de l'Embrapa, pour les soutenir dans des orientations à donner aux journalistes. Aussi sert-elle de dépôt de connaissance tacite, la transformant en explicite (Nonaka et Takeuchi, 1997) sur les spécialités et les compétences des chercheurs de l'Embrapa. Jusqu'à alors, ce type de connaissance était propriété exclusive de quelques chercheurs du Siège. Le Guide a été structuré sur base des informations suivantes : Nom, formation, secteur de recherche, thèmes/produits, unité de l'Embrapa, téléphone, courrier électronique et sujet pour la presse.

Penteado et Quoniam (2001, p.6) et Penteado et al. (2005b, p.16) ont détaillé les logiciels utilisés (Infotrans, Dataview, Matrisme, Microsoft Access, Microsoft Excel et Microsoft Word) et la logique du problème de recherche dans la gestion de compétences :

"Un des différentiels compétitifs réside, de plus en plus, dans la capacité d'administrer, d'une manière efficace la connaissance existante, soit dans le capital humain, soit dans les systèmes d'informations de l'organisation. Ceci signifie savoir identifier les connaissances exigées pour l'exécution des processus de gestion et de négoce et analyser les compétences existantes dans l'organisation, rendant possible de positionner les personnes correctement dans les postes de travail adéquats, dans le temps impartis, en transformant les informations en ressources d'intelligence dans un environnement compétitif"(Pitta, 2000).

En plus de l'addition de la base de R&D, la différence de la méthodologie des travaux de Penteado et al. (2005a, 2005b, 2003a et 2003b) a été l'union des Cohortes 1 et 2 (moins de 25 ans avec 25 à 34 ans) en fonction des effectifs très restreints :

- *Jonction du Guide de Sources et de la Base RH (Access). Résultat: une base avec 1791 chercheurs en étant 1609 actifs et 182 inactifs depuis 1997, avec pour chacun mention de leurs spécialités dans des Thèmes/Produits, local de travail et dates de naissance, début et fin du contrat ;*
- *Reformatage et normalisation (Infotrans) des références pour traitement sous le logiciel Dataview. Classification des dates de naissance. Création de quatre cohortes (1 - 25 à 34 ans ; 2 - 35 à 44 ans ; 3 - 45 à 54 ans ; 4 - plus de 55 ans) ;*
- *Analyse sur Dataview - traitement des champs cohorte, thèmes/produits, unité ;*
- *Génération de matrices de paires cohorte - thèmes/produits, cohorte - unité, thèmes/produits - unité ;*
- *Exportation sous Microsoft Excel des matrices ;*

- *Analyse des matrices - édition de graphiques et tableaux ;*
- *Exportation pour Matrisme des matrices carrées ;*
- *Analyse des matrices - édition de graphiques et tableaux (p.21).*

Penteado et al. (2005a et 2005b) ont étudié l'effectif de chercheurs concernant les compétences (figure 18) et ont produit une analyse plus détaillée des compétences des chercheurs plus jeunes (figure 19), les compétences des chercheurs seniors que sont en danger d'être perdues par l'organisation (figure 20), outre les enquêtes des effectifs par des unités de recherche (figures 21, 22), des tableaux des compétences par les différentes cohortes (figure 23) et les compétences perdues, des chercheurs qui pour des diverses raisons (décès, retraite, démission, autres) ont quitté l'Embrapa (figure 24). L'analyse considère 1791 chercheurs du total de 2221 chercheurs de l'Embrapa. On peut considérer que les 430 manquants (19,3%) sont des champs vides à combler.

Parmi 1.791 chercheurs de l'échantillon, 53 (3%) ont de 25 à 34 ans, 501 (28%) ont de 35 à 44 ans, 806 (45%) ont de 45 à 54 ans et 430 (24%) ont plus de 55 ans. En additionnant les deux dernières cohortes, 69% des chercheurs ont plus de 45 ans comme le présente la figure 18.

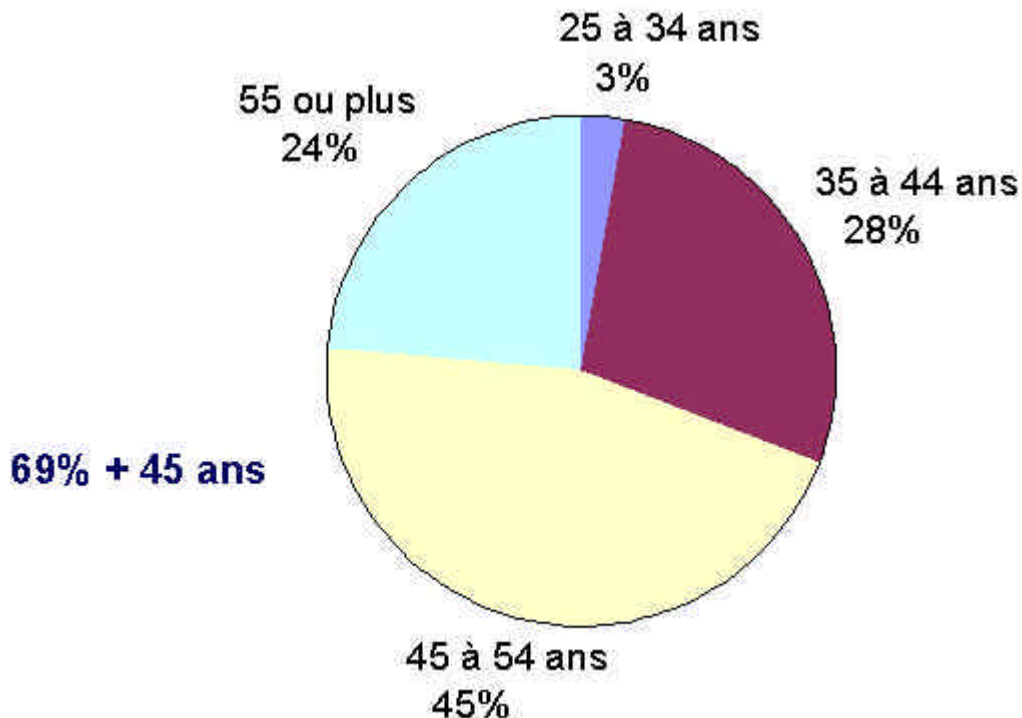


Figure 18 : Distribution des chercheurs par cohorte (Penteado et al., 2005a, p.24)

Des 632 compétences analysées dans le Guide de Sources, 14 ont un chercheur dans la cohorte jusqu'à 30 ans et 618 n'ont personne. Tous les chercheurs occupent les cohortes ultérieures. Comme ce dernier nombre représente 97% des compétences, c'est un signe d'alerte sur la nécessité de renouvellement des équipes, comme l'indique la figure 19. Les compétences marquées avec des étoiles : Informatique, biotechnologie, banane, bio fertilisants et acerola ont relativement peu de chercheurs dans la dernière cohorte, néanmoins, existent des compétences qui même en ayant quelques jeunes chercheurs inspirent plus d'attention, identifiées par des flèches : Arboriculture fruitière (**fruticultura**), fertilité du sol, consortiums de plantes, pâturages et fourragères (**pastos e forrageiras**), sols/conservation et manège comme montré dans la figure 19.

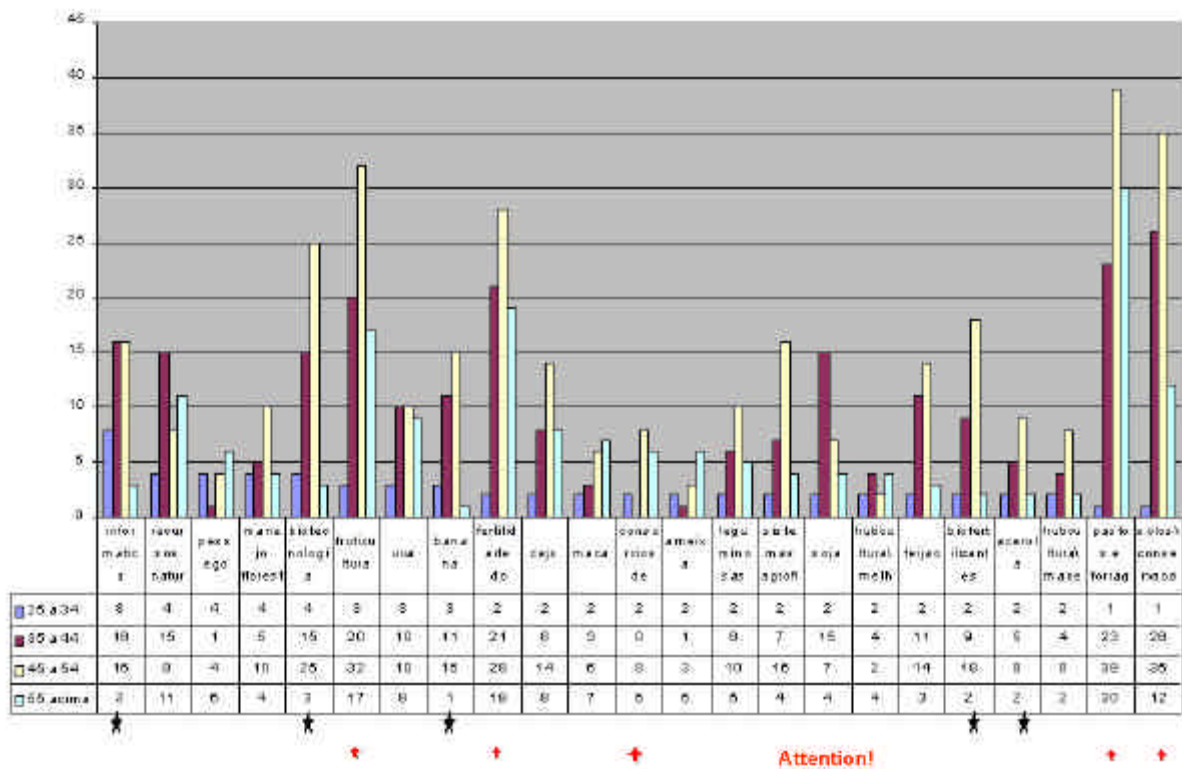


Figure 19 : Compétences des plus jeunes (Penteado et al., 2005a, p.25)

La figure 20 indique les spécialités qui concentrent les plus grands effectifs de chercheurs avec plus de 55 ans. De ces spécialités, quelques-unes, comme sols/classement et enquête (**solos/classificacao e levantamento**), zonage (**zoneamento**), production horticoles (**hortalicas**), ressources naturelles (**recursos naturais**), bovines/pâturages (**bovinos/pastagens**), riz (**arroz**) et raisin (**uva**) comptent avec des effectifs significatifs dans la cohorte de 35 à 44 ans, ce qui peut indiquer une moindre urgence dans l'embauche de

nouveaux chercheurs. D'autres spécialités appellent l'attention, comme celles identifiées avec des flèches : Pâturages et fourragères (**pastos e forrageiras**), arboriculture fruitière (**fruticultura**), fertilité du sol (**fertilidade do solo**), sols\conservation et manège. D'autres désignées avec des flèches plus grosses, comme économie agricole, systèmes de production, bovins/lait, semences, Palmier à huile (**dende**), contrôle de pestes (**contrôle de pragas**), bovins\nutrition et soja/manège et traitements de la culture, outre d'avoir entre 30 et 9 chercheurs de plus de 55 ans, ont des effectifs de la cohorte suivante (de 45 à 54 ans) deux fois ou plus petit de ce de la cohorte ultérieure, de 35 à 45 ans. Ce tableau présenté figure 20 indique les compétences où il y a la plus grande nécessité d'embaucher de nouveaux chercheurs.

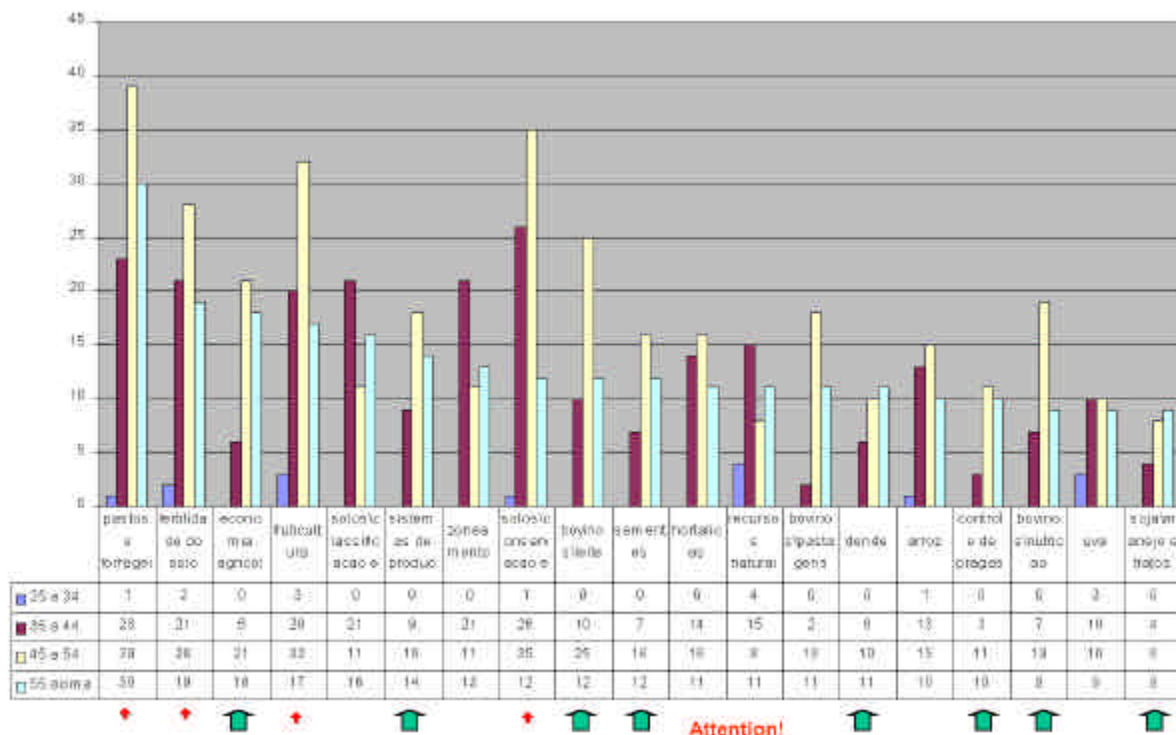


Figure 20 : Compétences en risque (Penteado et al., 2005a, p.26)

Les figures suivantes répondent aux questions de savoir où sont les plus jeunes et où sont les plus vieux chercheurs. Les deux tableaux permettent des observations sur la nécessité de renouvellement dans des centres de recherche spécifiques. Figure 21 indique les centres avec les plus grands effectifs de chercheurs avec 25 à 34 ans, marqués avec une étoile. Ce sont, Informatique Agricole (**Informatica Agropecuaria**), Raisin et Vin (**Uva e Vinho**), Acre et Roraima. Ces quatre centres de recherche ont de huit à quatre chercheurs dans cette cohorte.

En outre dans cinq unités, Amazonie Occidentale (**Amazonia Ocidental**), Porcs et Volailles (**Suinos e Aves**), Pantanal et Amapa les chercheurs jeunes sont majoritaires, en ajoutant les deux cohortes de 25 à 34 ans et de 35 à 44 ans. Ces centres sont aussi marqués par des étoiles. Dans six centres de recherche, désignés avec des flèches, le nombre de chercheurs avec plus de 45 ans dépasse le plus du double l'effectif des chercheurs les plus jeunes : Amazonie Orientale, Siège, Semi-aride, Cerrados, Riz et Haricot et Climat Tempéré. Tous ces résultats sont présentés figure 21.

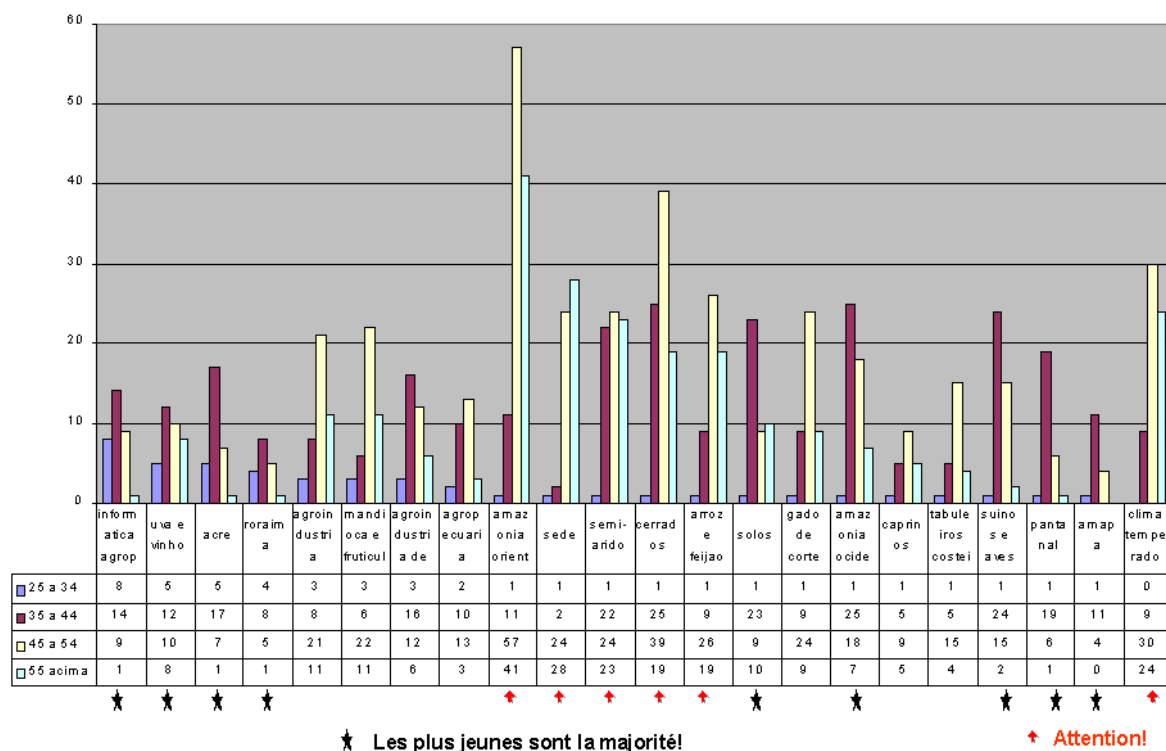


Figure 21 : Où sont les plus jeunes (Penteado et al., 2005a, p.27)

Figure 22 indique les unités qui ont besoin d'attention en raison des pertes possibles de compétences. Six ont été identifiées précédemment et sont marquées avec des étoiles : Amazonie Orientale, Siège (**Sede**), Semi-aride, Cerrados, Riz et Haricot (**Arroz e Feijao**) et Climat Tempéré (**Clima Temperado**). Dix autres centres de recherche ont plus de deux tiers des leurs chercheurs âgés de plus de 45 ans : Maïs et Sorgho (**Milho e Sorgo**), Soja, Blé (**Trigo**), Moyen-Nord (**Meio-Norte**), Agro-industrie Tropicale, Manioc et Arboriculture Fruitière Tropicale (**Mandioca e Fruticultura Tropical**), Bovins de Lait (**Gado de Leite**), Bovins de Viande (**Gado de Corte**), Coton (**Algodao**) et Élevage du Bétail-Sud (**Pecuaria**

Sul). Penteadó et al. (2003) ont remarqué qu'il est naturel d'en trouver, au siège des organisations, un plus grand nombre de personnes âgées. Plus de détails dans la figure 22.

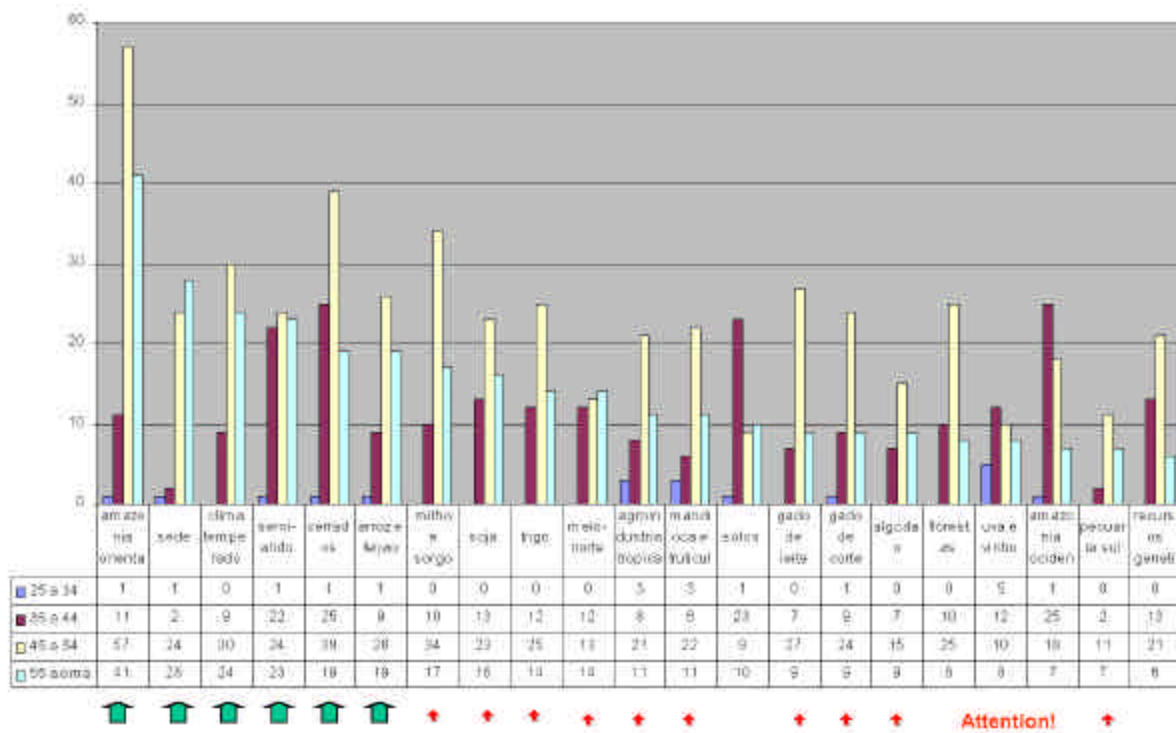


Figure 22 : Où sont les plus vieux (Penteadó et al., 2005a, p.28)

En travaillant avec les grands thèmes de recherche, il est possible de comparer les 20 compétences avec la plus grande fréquence des cohortes 1 (plus jeunes) et 4 (plus vieux). Les résultats indiquent des compétences qui restent, qui apparaissent et qui disparaissent. Ces dernières : Contrôle intégré de pestes et des maladies (**Controle integrad. pragas e doencas**), amélioration végétale (**Melhoramento vegetal**), socioéconomie (**Socioeconomia**), transfert et diffusion de technologie (**Transferencia e difusao de tecnologia**), fourrages et nutrition animale (**Forragicultura e nutricao animal**), physiologie végétale (**Fisiologia vegetal**) et après-cueillette (**Poscolheita**), sont signalées avec des étoiles dans la figure 23. Elles indiquent des compétences où on ne se produit plus de renouvellement.

Champs qui continuent

	T	C1
Fitotecnia	265	5
Manejo conserv. solo / nul plant	226	2
Recursos Genéticos	87	6

Champs qui disparaissent

	T	C1
Controle integrad pragas e doenças	227	1★
Melhoramento vegetal	169	2★
Socioeconomia	124	0★
Transferência e difusão tecnologia	112	1★
Forragicultura e nutrição animal	91	1★
Fisiologia vegetal	71	1★
Poscolheita	108	1★

Champs qui apparaissent

	T	C1
Manejo animal	68	9
Recursos naturais	74	9
Sistemas florestais	74	6
Monitoramento ambiental	93	5
Biologia avançada	97	4
Métodos quantitativos	32	3
Sistemas irrigados	49	3

Figure 23 : Évolution des compétences par secteur de recherche Cohortes 1 et 4 (Penteado et al., 2005a, p.33)

Pour confirmer les résultats sur les compétences dans des secteurs de recherche qui disparaissent, la même comparaison a été réalisée concernant les 20 compétences avec les plus grandes fréquences dans les cohortes 2 (de 35 à 44 ans) et 4 (plus vieux - plus de 55 ans). Cette fois, les sept secteurs de recherche qui ont disparu restent. Mais seulement l'un d'eux, l'après cueillette (**Poscolheita**), présente un haut taux de renouvellement de 38%. Il est marqué avec une étoile de quatre branches dans la figure 24, ainsi que le secteur de biologie avancée (**Biologia avancada**) qui a le plus haut taux de renouvellement, de l'ordre de 42%. Quatre secteurs de recherche : Contrôle intégré de pestes et maladies (**Controle integrado prag/doenc**), amélioration végétale (**Melhoramento vegetal**), fourrages et nutrition animale (**Forragicultura e nutricao animal**), physiologie végétale (**Fisiologia vegetal**), marqués avec des étoiles de cinq branches dans la figure 24, ont des taux de renouvellement entre 12,6% et 14%. Deux secteurs de recherche, socioéconomie (**Socioeconomia**) et transfert et diffusion de technologie (**Transferencia e difusao de tecnologia**), indiqués avec des flèches dans la figure 24, ont des taux de renouvellement de, respectivement, 9,6% et 8,9%. Ces six secteurs méritent l'accompagnement quant au renouvellement de leurs effectifs. Parmi les secteurs qui disparaissent, méthodes quantitatives, marqué avec une flèche plus grosse, a un taux de

renouvellement inférieur à 10%. A remarquer le surgissement du secteur de l'agriculture de précision (**Agricultura de precisao**) comme illustré figure 24.

Champs qui continuent			Champs qui apparaissent		
	T	C2		T	C2
Fitotecnia	265	44	Agricultura de precisão	29	7↑
Controle integrado prag/doenc	227	34★	Outras áreas tecnológicas	40	24
Manejo e conserv. solo /nut plant	226	41			
Melhoramento vegetal	169	24★			
Socioeconomia	124	12★			
Transferência/difusão tecnologia	112	10★			
Forragicultura e nutrição animal	91	12★			
Fisiologia vegetal	71	9★			
Poscolheita	108	42↑			
Manejo animal	68	17			
Sanidade animal	49	7	Métodos quantitativos	32	3←
Monitoramento ambiental	93	22	Melhoramento animal	30	12
Sistemas florestais	74	19			
Sistemas irrigados	49	7			
Recursos naturais	74	26			
Biologia avançada	97	41↑			
Reprodução animal	30	12			

Figure 24 : Évolution des compétences par secteur de recherche Cohortes 2 et 4 (Penteado et al., 2005a, p.34)

Quand nous étudions les compétences perdues des 72 chercheurs qui ont laissé l'Embrapa pour des diverses raisons entre juin 1996 et juillet 2001, trois secteurs indiqués ci-dessus pour accompagnement réapparaissent : Socioéconomie (**Socioeconomia**), transfert et diffusion de technologie (**Transferencia de tecnologia**) amélioration végétale (**Melhoramento vegetal**). Ils sont identifiés avec des étoiles dans la figure 25. Comme le secteur de socioéconomie a été celui qui a perdu le plus de chercheurs, neuf au total, il a gagné deux étoiles dans la figure 25.

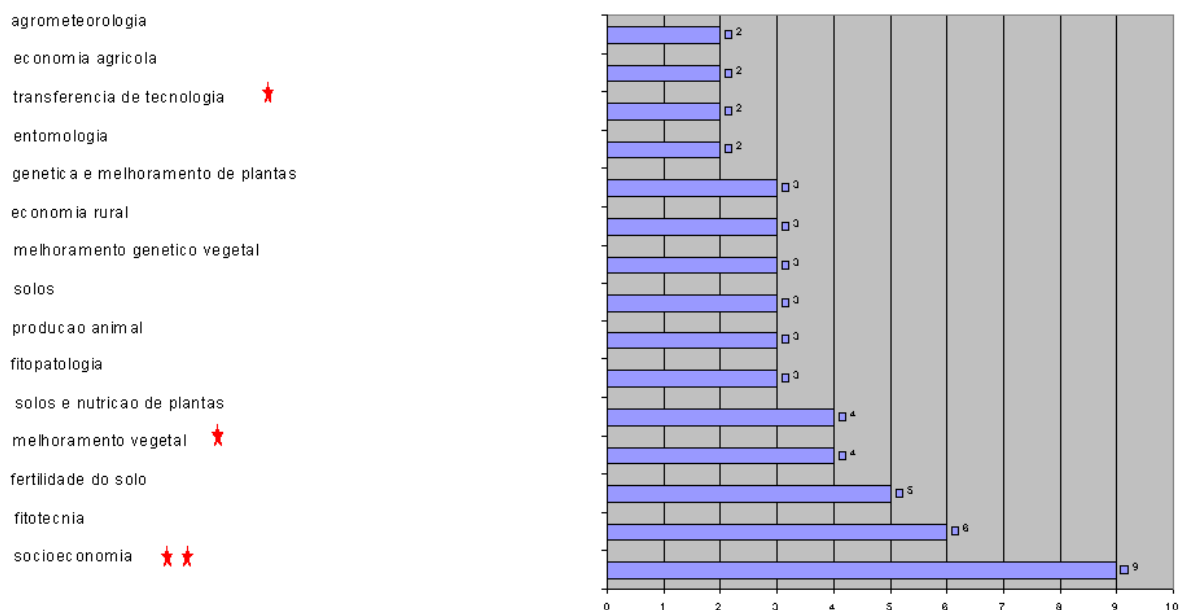


Figure 25 : Compétences perdues (Penteado et al., 2005a, p.35)

La croissance des sorties en 2005

Une mise à jour de la base de compétences des chercheurs de l'Embrapa, en juillet 2005, a révélé une augmentation considérable de départ de chercheurs après 2002. La figure 26 rapporte les départs de chercheurs depuis 1996, avec un total de 303 chercheurs.

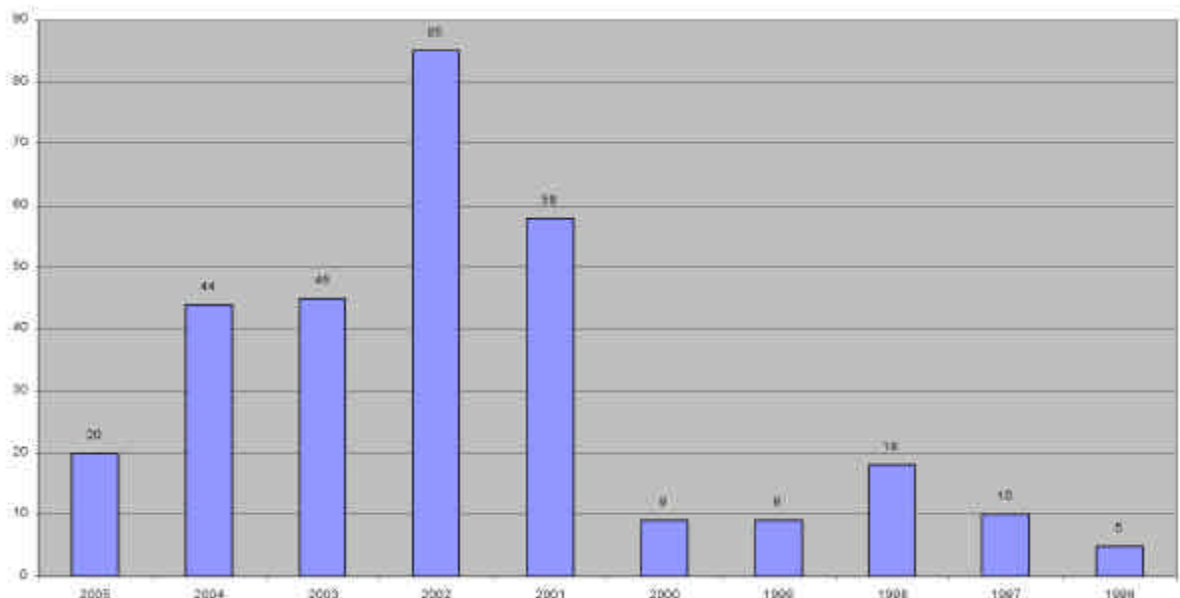


Figure 26 : 303 Chercheurs sortants depuis 1996

Des six secteurs indiqués pour suivre leur renouvellement : Amélioration végétale (**Melhoramento vegetal**), transfert et diffusion de technologie (**Transferencia e difusao de tecnologia**), socioéconomie (**Socioeconomia**), fourrages et nutrition animale (**Forragicultura e nutricao animal**), après-cueillette_transf.agro-industrielle (**Pos-colheita_transf. agroindustrial**) et contrôle intégré de pestes et des maladies (**Contrôle integrado de pragas e doencas**), quatre sont entre les huit secteurs qui ont perdu le plus d'employés entre 1996 et 2005. L'un d'eux a été le secteur qui a plus perdu des ressources humaines, *transfert et diffusion de technologie*, avec 22 chercheurs, ce qui a représenté 26,8% de l'effectif total. Le secteur de socioéconomie a perdu 21 chercheurs ou 16,1% de son effectif total. Sept autres secteurs aussi ont perdu plus de 15% de leurs effectifs : Phytotechnique (**Fitotecnica**, 21,3%), manège et conservation du sol et nutrition de plantes (17,3%), amélioration végétale (15,6%), fourrages et nutrition animale (17,4%), manège animal (**Manejo animal**, 15,6%), méthodes quantitatives (**Metodos quantitativos**, 15,8%) et santé animale (**Sanidade animal**, 15,4%). Les employés qui ont quitté l'entreprise appartenaient à tous les grands secteurs de recherche comme le montre les détails du tableau 51.

Tableau 51 : Compétences perdues par secteurs de recherche

Secteurs de recherche	Chercheurs Sortants	Total Chercheurs	%
Fitotecnia	51	239	21,3
Manejo e conserv. solo e nutricao de plantas	43	248	17,3
Melhoramento vegetal	29	186	15,6
Transferencia e difusao de tecnologia	22	82	26,8
Controle integrado de pragas e doencas	22	232	9,5
Socioeconomia	21	130	16,1
Forragicultura e nutricao animal	16	92	17,4
Pos-colheita_ transf. agroindustrial	12	113	10,6
Manejo animal	12	77	15,6
Recursos naturais	9	73	12,3
Monitoramento ambiental	9	101	8,9
Recursos geneticos	9	81	11,1
Sanidade animal	8	52	15,4
Sistemas irrigados	6	56	10,7
Metodos quantitativos	6	38	15,8
Sistemas florestais	5	77	7,1
Reproducao animal	4	31	12,9
Melhoramento animal	3	22	13,6
Fisiologia vegetal	3	78	3,8
Agricultura de precisao	3	28	7,1
Total	293	2036	14,4

Quand nous ajoutons les embauches de nouveaux chercheurs entre 2002 et 2005 notre univers d'analyse s'agrandit en 2005 pour 2231 chercheurs. En considérant 27 secteurs de recherche, le nombre de chercheurs de 55 ans ou plus revient à un pourcentage relativement élevé, 31% (691). Onze secteurs ont des effectifs au-dessus de la moyenne. Ils sont marqués en gris dans le tableau 52 : Transfert et diffusion de technologie (**Transferencia e difusao de tecnologia**),

d'autres champs technologiques (**Outras areas tecnologicas**), socioéconomie (**Socioeconomia**), fourrages et nutrition animale (**Forragicultura e nutricao animal**), phytotechnique (**Fitotecnia**), contrôle intégré de pestes et maladies (**Controle integrado de pragas e doencas**), amélioration végétale (**Melhoramento vegetal**), méthodes quantitatives (**Metodos quantitativos**), manège et conservation du sol et nutrition des plantes (**Manejo e conservacao do solo e nutricao de plantas**), santé animale (**Sanidade animal**) et amélioration animale (**Melhoramento animal**). Si on fait référence aux compétences perdues, des six secteurs contrôlés, il n'y a que dans celui d'après-cueillette_transf. agro-industrielle (**Pos-colheita _transf. agroindustrial**), avec 15,9%, le pourcentage de chercheurs de 55 ans ou plus reste au-dessous de la moyenne. Le premier secteur contrôlé, *transfert et diffusion de technologie* enregistre un pourcentage presque double de la moyenne, 59,8%. Tous les autres secteurs contrôlés ont aussi des effectifs de chercheurs avec 55 ans ou plus au-dessus de la moyenne : Socioéconomie (44,6%), fourrages et nutrition animale (44,6%), contrôle intégré de pestes et maladies (35,8%), amélioration végétale (34,9%), méthodes quantitatives (34,2%). Le tableau 52 montre la situation des secteurs de recherche des chercheurs avec 55 ans ou plus.

Tableau 52 : Secteurs de recherche des chercheurs avec 55 ans ou plus

Secteurs de recherche	55 ans ou plus	Total Chercheurs	%
Transferencia e difusao de tecnologia	49	82	59,8
Outras areas tecnologicas	6	13	46,2
Socioeconomia	58	130	44,6
Forragicultura e nutricao animal	41	92	44,6
Fitotecnia	102	239	42,7
Controle integrado de pragas e doencas	83	232	35,8
Melhoramento vegetal	65	186	34,9
Metodos quantitativos	13	38	34,2
Manejo e conserv. solo e nutricao de plantas	81	248	32,7
Sanidade animal	17	52	32,7
Melhoramento animal	7	22	31,8
Fisiologia vegetal	24	78	30,8
Analise instrumental	3	10	30
Recursos geneticos	23	81	28,4
Aqüicultura	3	11	27,3
Manejo animal	21	77	27,3
Reproducao animal	8	31	25,8
Sistemas irrigados	13	56	23,2
Sistemas florestais	16	77	20,8
Pos-colheita_ transf. agroindustrial	18	113	15,9
Monitoramento ambiental	14	101	13,9
Recursos naturais	10	73	13,7
Biologia avancada	12	111	10,8
Recursos da fauna	1	12	8,3
Agricultura de precisao	2	28	7,1
Modelagem e simulacao de sistemas	1	30	3,3
Campovazio	0	5	0
Apicultura	0	3	0
Total	691	2231	31

Si l'âge du chercheur senior est réduit à 50 ans ou plus, le nombre de chercheurs concernés monte de 691 pour 1042 ou 46,7% du total de 2231. Neuf secteurs de recherche parmi les 27 restent avec des effectifs de 50% ou plus de chercheurs seniors avec plus de 50 années : Transfert et diffusion de technologie, socioéconomie, fourrages et nutrition animale, phythotechnique, amélioration animale, méthodes quantitatives, contrôle intégré de pestes et maladies, physiologie végétale et amélioration végétale. Six des sept secteurs contrôlés quant à la perte de compétences y sont rapportés.

Quand nous analysons les 25 lignes de recherche liées aux onze secteurs de recherche déjà identifiés avec des effectifs au-dessus de la moyenne, un résultat pratique est l'identification des spécialités des chercheurs à engager, leurs quantités et leurs régions. Parmi 15 lignes de recherche considérées prioritaires par l'entreprise, sept d'elles sont entre les dix avec la plus grande proportion de chercheurs de 55 ans ou plus. À noter que les deux lignes de recherche du secteur le plus critique, transfert et diffusion de technologie, sont aussi les deux lignes de recherche avec des pourcentages de chercheurs de 55 ans ou plus de, respectivement, 81,8% et 52,5%.

Les effectifs de treize lignes de recherche ont un nombre de chercheurs plus vétérans au-dessus de la moyenne de 38,5%. Elles sont (l'astérisque et les mentions nac/no/ne/co/su/s indiquent qui sont des lignes prioritaires nationales ou régionales, nac=national, no=nord, ne=nord-est, co=centre-ouest, se=sud-est, s=sud) : *(no_nac) transfert et diffusion de technologie (**Transferencia e difusao de tecnologia**), *(ne_co_nac) communication et négoce (**Comunicacao e negocios**), *(no) forragiculture (**Forragicultura**), *(no) technologie en bois (**Tecnologia de madeiras**), *(se_s_nac) environnement et constructions agricoles (**Ambiencia e construcoes rurais**), économie économie agricole ressources naturelles (**Economia economia rural recursos naturais**), *(s) fertilité du sol (**Fertilidade do solo**), *(co) technologie de semences et boutures (**Tecnologia de sementes e mudas**), phytopathologie (**Fitopatologia**), systèmes de production (mécanisation agricole plantation directe_production organique) [**Sistemas de producao (mecanizacao agricola_plantio sireto_producao organica)**], *(se) nutrition de ruminants (**Nutricao de ruminantes**), amélioration de cultures pérennes (**Melhoramento de culturas perenes**), *(se) pathologie_y compris immunologie et parasitologie (**Patologia incluindo imunologia e parasitologia**). Elles sont marquées en gris au tableau 52. En considérant la moyenne totale de chercheurs avec âge de 55 ans ou plus de l'entreprise, 31%, s'ajoutent autres six lignes de recherche :

*(no_ne) genèse_morphologie et physique du sol (**Genese_morfologia e fisica do solo**), *(se) amélioration génétique animale (**Melhoramento genetico animal**), amélioration de cultures annuelles (**Melhoramento de culturas anuais**), sociologie_sociologie agricole (**Sociologia sociologia rural**), méthodes quantitatives dans R&D (**Metodos quantitativos em p&d**), entomologie (**Entomologia**). Le tableau 53 présente les détails.

Tableau 53 : Lignes de recherche avec plus de chercheurs seniors

Lignes de recherche (*=priorité nationale ou régionale)	55 ans ou plus	Total Chercheurs	%
*(no_nac) Transferencia e difusao de tecnologia	18	22	81,8
*(ne_co_nac) Comunicacao e negocios	28	53	52,8
*(no) Forragicultura	27	53	50,9
*(no) Tecnologia de madeiras	3	6	50
*(se_s_nac) Ambiencia e construcoes rurais	2	4	50
Economia_ economia rural_ recursos naturais	51	107	47,7
*(s) Fertilidade do solo	36	81	44,4
*(co) Tecnologia de sementes e mudas	24	55	43,6
Fitopatologia	53	123	43,1
Sistemas de producao (mecanizacao agricola_ plantio direto_ producao organica)	79	186	42,5
*(se) Nutricao de ruminantes	13	31	41,9
Melhoramento de culturas perenes	24	58	41,4
*(se) Patologia_ incluindo imunologia e parasitologia	14	35	40
*(no_ne) Genese_ morfologia e fisica do solo	16	42	38
*(se) Melhoramento genetico animal	7	20	35
Melhoramento de culturas anuais	41	127	32,3
Sociologia_ sociologia rural	7	22	31,8
Metodos quantitativos em p&d	13	41	31,7
Entomologia	24	77	31,2
Nutricao de monogastricos	3	11	27,3
Manejo e conservacao do solo	21	78	26,9
*(co) Agricultura conservacionista	1	4	25
*(co_s) Microbiologia do solo	7	29	24,1
*(co) Epidemiologia_ incluindo bacteriologia e virologia	3	17	17,6
*(co_nac) Controle biologico	6	37	16,2
Total 1 (Chercheurs des 25 Lignes de recherche étudiés)	521	1319	38,5
Outres lignes de recherche	170	912	18,6
Total 2 (Total de chercheurs de l'Embrapa)	691	2231	31

Le développement de tableaux et graphiques pour toutes les lignes de recherches de l'organisation permet d'identifier avec une raisonnable précision, pour chaque sujet, produit, ligne, thème ou secteur scientifique, les nécessités d'embauche de nouveaux chercheurs.

Conclusion

Cet étude est le résultat de l'intégration d'informations de trois bases de données, des ressources humains (SIRH), de l'agenda de R&D et du guide de sources, de trois départements différents de l'Embrapa, respectivement, le Département de Gestion de Personnes (DGP), la Surintendance de Recherche et Développement (SPD) et le Bureau de Communication Sociale (ACS).

Il démontre comme ces informations déjà disponibles dans beaucoup d'organisations, mais dispersées en diverses places, peuvent devenir utiles pour l'action, soit pour le transfert et la diffusion de technologie, soit pour l'administration stratégique de compétences. Il fait ressortir aussi que les organisations ne peuvent plus ignorer ce fait : Il y a beaucoup d'informations en leur sein qui sont très peu utilisées. Comme expliquent Penteado et al. (2003, p.6) "stocker des données pour une utilisation faisant simplement appel à 10 ou 20% des utilisations possibles conduit à une perte de temps et aussi à un affaiblissement stratégique".

Les méthodologies et outils ici démontrées permettent de générer de l'intelligence avec l'agilité nécessaire pour la prise de décisions et l'action. Elles indiquent aussi une voie pour qui les organisations initient la construction et la gestion de la fonction d'analyse : Disséminer de telles analyses parmi les différents secteurs comme les ressources humaines, matériels, financiers, recherche et développement, entre autres. Quand de telles analyses et rapports d'intelligence deviennent institutionnels et leurs informations, des indicateurs de gestion de l'intelligence, l'organisation sera plus proche de la création d'un système d'informations stratégiques.

Tout se passe comme décrit Van Raan (1997, apud. Faria, 2001, p.22) : Quand il y a des études focalisées et d'une durée relativement courte, comme des travaux d'intelligence économique, la préparation des données normalement est développée spécifiquement pour cette tâche. Pour des études systématiques et de durée plus longue, comme la production d'indicateurs pour l'évaluation de la production scientifique, la préparation des données devient un processus continu et systématique, en visant à la création d'un système

d'information pérenne, amélioré au fil du temps avec l'expérience résultante d'applications successives.

Autre point à remarquer c'est que la création d'un environnement informatique spécifique pour l'analyse, en réunissant toutes les bases opérationnelles et des informations d'origines diverses, internes et externes, augmente la productivité des analystes et rend possible des études dans des délais plus rapides.

Analyse de la performance de l'Embrapa dans les médias

Le système de Clipping électronique de l'Embrapa

Créé en 1997, le système de Clipping électronique de l'Embrapa, la veille sur les nouvelles sorties dans les médias concernant l'Entreprise, a introduit des concepts innovateurs, en utilisant des technologies comme l'Internet et l'Adobe Acrobat pour digitaliser tout ce qui était fait précédemment sur papier et offrir ce nouveau produit par un réseau interne, l'Intranet de l'Embrapa, aux 40 centres de recherche. Une méthodologie simple et efficace pour produire un instrument important pour la démocratisation des informations dans les organisations. Le clipping est défini comme la relation du texte complet des nouvelles sorties dans les médias sur l'organisation. Il est distribué une fois par jour, le matin, dans la majorité des cas.

Bueno (2002) affirme que :

Incomplet ou pas, inexact ou pas, confondu ou pas, le clipping est la (...) matière première fondamentale d'un travail d'audit, à être fait a posteriori, que, si bien conduit, pourra signaler des opportunités de divulgation, diagnostiquer des personnalités et styles des véhicules et des éditeurs et surtout, permettre que les entreprises ou les organisations affinent leur travail de relations avec les mass médias.

La veille du macro-ambient est une priorité et une nécessité des organisations. L'objectif est d'accompagner, en premier lieu, ce qui sort sur l'organisation dans les médias et en deuxième place, ce qui sort sur leurs concurrents. Peu d'organisations accomplissent ce second objectif parce que la majorité absolue n'arrive pas à faire le premier correctement.

Le clipping sous format papier a des graves limitations de dimension et de prix pour le distribuer à plus d'une dizaine de personnes. Ce sont des dizaines de milliers de feuilles photocopiées par mois, ce qui rend difficile et laborieux le maintien de l'archive. Ainsi, seulement la direction ou la haute couche d'une organisation a l'occasion d'être informée sur ce qui se passe en son macro-ambient. L'employé commun et les gérants et les responsables restent sans accès aux informations cruciales pour leur quotidien, leurs projets et initiatives.

Comme l'ancien système de clipping était tout en papier, l'économie avec des photocopies a été aussi significative. Dans les six premiers mois d'opération du clipping électronique, de juillet à décembre 1997, le nombre de photocopies dans le Bureau de Communication Sociale a décré de 29.854 à 15.086, une réduction de 49,46%.

Avec l'impulsion du sous projet 1. Création d'un système pilote d'intelligence compétitive et de gestion de la connaissance pour la communication à l'Embrapa, nous avons cherché à tracer quelques directives pour la création d'une banque de données intelligentes sur le matériel réuni dans la base Clipping.

Le Clipping électronique a évolué, en 2002, du HTML pour le Lotus Notes. L'utilisateur peut accéder la banque de données en sélectionnant un véhicule (plus de 962 journaux ou des revues), une date (Jour/Mois/An), un centre de recherche (39 Unités) ou un thème ou produit (84 au total).

Objectifs

Le travail d'analyse de la performance de l'Embrapa et de ses centres de recherche dans les mass médias avec la base Clipping vise à mieux qualifier et ajouter de l'intelligence au travail réalisé dans les secteurs de communication des centres de recherche. Ces objectifs sont:

- Identifier la politique éditoriale des journaux en contribuant par cela à qualifier et améliorer le travail de divulgation dans les mass médias, en augmentant cette couverture.
- Automatiser l'analyse et la génération d'indicateurs d'audit des médias et éventuellement de substituer le travail actuellement acheté des tierces parties "d'analyse quantitative et qualitative des citations de l'Embrapa dans des journaux et revues".
- Rendre disponible l'intelligence aux gérants et aux journalistes des centres de recherche, avec l'utilisation des outils appropriés d'analyse de données et de textes, la réalisation de leurs propres analyses et explorations des données du Clipping de leur centre de recherche.
- Faire l'audit des travaux de communication sociale réalisés par l'Embrapa en ses centres de recherche. Le Bureau de Communication Sociale de l'Embrapa (ACS) et les bureaux de communication des centres de recherche travaillent d'une manière intégrée : L'ACS reçoit des suggestions de nouvelles des centres de recherche et travaille avec ces suggestions dans une divulgation nationale. Les centres de recherche travaillent dans la divulgation régionale, dans les états (départements) et locale (villes). Ce système de travail doit aussi être évalué.

Méthodologie

La base Clipping

Originellement, la base Clipping avait les champs date du clipping, titre, véhicule, date de la publication, thème ou produit, page, centre de recherche et mots-clés.

Dans la mise à jour pour le logiciel Lotus Notes, de nouvelles variables ont été créés. Voyez le tableau 54.

Tableau 54 : Les nouvelles variables de la base Clipping

- | |
|---|
| <ol style="list-style-type: none">1. Titre de la nouvelle (TIT) ;2. Date de la nouvelle (DTP) ;3. Date de clipping (DCLIP) ;4. Nom du Véhicule (VEI) ;5. Département (État) du Véhicule (UF) ;6. Sujet ou Thème de la nouvelle (classé selon une liste de mots-clés contrôlés par l'Embrapa) (TEM) ;7. Genre journalistique de chaque nouvelle (texte, éditorial, reportage, interview, article, note d'opinion, note d'information, lettre du lecteur, chronique) (GEN) ;8. Type de la présence de l'organisation dans la nouvelle (cape/1^{ère} page, titre de page, titre, prééminence dans le texte, citation, plinthe/légende) (PRE) ;9. Source mentionnée dans la nouvelle (dirigeante, chef de centre, chercheur, autres employés, source non cité) (FON) ;10. Page de publication (pair, impair, deux pages, trois pages, quatre ou plus pages) (PAG) ;11. Localisation de la nouvelle dans la page (supérieur droit, supérieur gauche, inférieur droit, inférieur gauche, toute la page) (QUA) ;12. Traitement graphique, nombre d'éléments présents (un élément - texte, deux éléments - texte et photo ou illustration, trois éléments - texte, photo/illustration et box, quatre éléments - texte, photo, illustration et box, cinq ou plusieurs éléments) (CGR) ;13. Unité(s) de l'Embrapa mentionnée(s) (UD) ;14. Mots-clés de la nouvelle (PCH) ;15. Positionnement de la nouvelle concernant la mission de l'Embrapa (POS*) ;16. Polarisation des nouvelles (POL*) ;17. Liaison pour l'archive digitalisée (LNK). |
|---|

Note. * Les champs indiqués avec l'astérisque n'ont pas été informés dans la base Clipping.

Voyez dans la figure 27, un exemple d'une page du clipping avant la digitalisation et dans la figure 28 cette même page après son entrée au logiciel Lotus Notes.



Figure 27 : Page du Clipping Électronique avant digitalisation

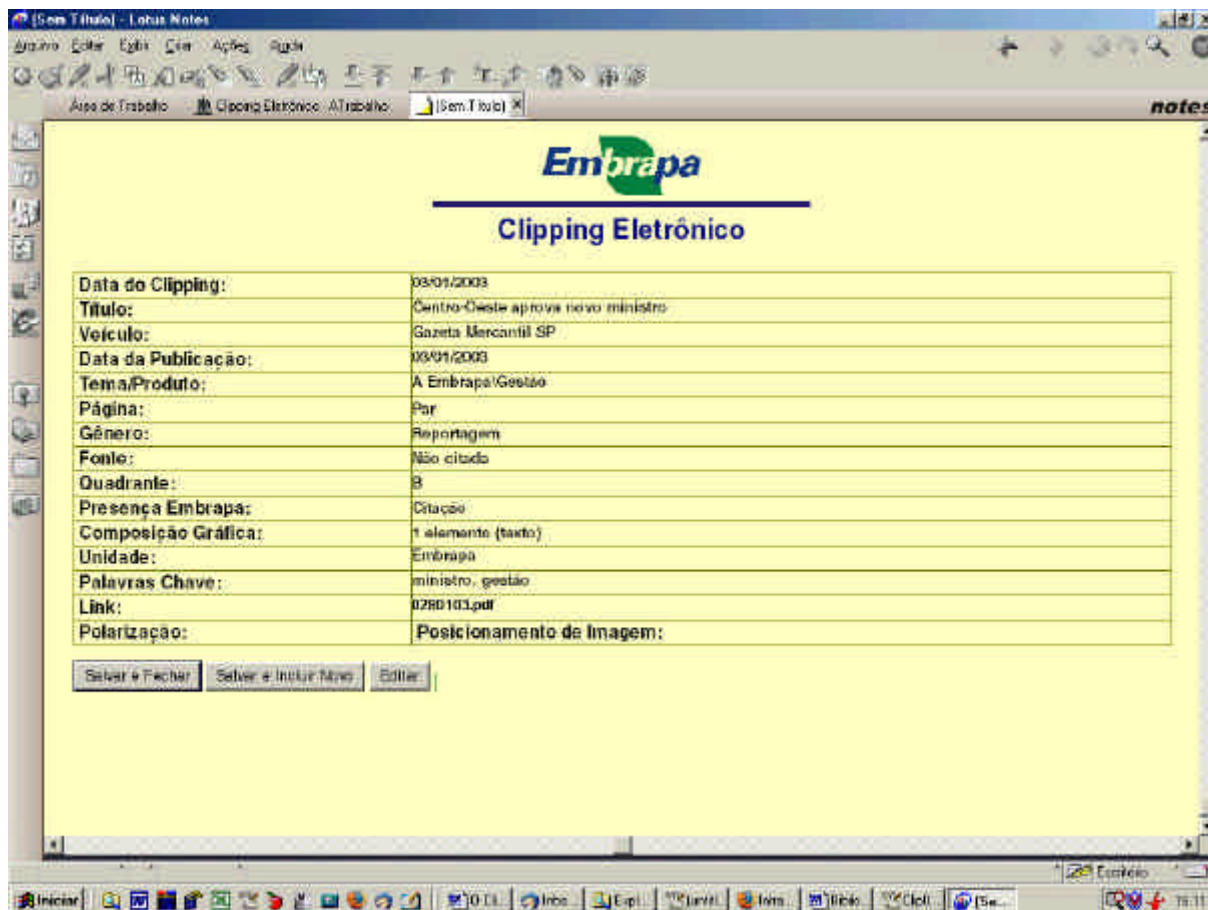


Figure 28 : Entrée de la page digitalisée dans le Lotus Notes

Les références du clipping des années de 2002, 2003, 2004 et 2005 ont été exportées de Lotus Notes et reformatés dans l'Infotrans. Au total il y avait 17 champs et 23.765 références. La base reformatée dans l'Infotrans a été importée dans Microsoft Access. Un audit a été réalisé et 189 références ont été identifiées inconsistantes ou vides. Ils ont été enlevés. Les champs 3. Date du clipping (DCLIP), 17. Liaison pour l'archive digitalisée (LNK) et les deux champs non informés dans la base, 15. Positionnement de la nouvelle concernant la mission de l'Embrapa (POS) et 16. Polarisation des nouvelles (POL) n'ont pas été considérés pour l'analyse. Ensuite, la base avec 13 champs a été exportée pour texte ASCII en format .CSV concernant les clippings des années 2003 et 2004, dans un total de 16.999 références. La nouvelle base a été reformatée dans Infotrans. En résultent des références comme celle présentée tableau 55.

Tableau 55 : La base Clipping reformatée

TIT : Na ponta dos dedos
VEI : Panorama Rural
UF : SP
DTP : 01/2003
ANO : 2003
PAG : 4 ou + paginas
GEN : Noticia
FON : Pesquisador
PRE : Destaque no texto
CGR : 2 elementos (texto + foto/ilustracao)
TEM : Fatores e Insumos, Maquinas e Implementos
PCH : agricultura
QUA : A, B, C, D, E
UD : Milho e Sorgo, Instrumentacao Agropecuaria

Véhicules

Les véhicules ont été segmentés selon leur distribution : Journaux nationaux - vendus dans tout le territoire national : Principaux journaux – avec influence régionale, Journaux des états – avec influence dans l'état, Agronégoce – spécialisés en ce secteur, Revues – circulation nationale, Bulletins – circulation restreinte, Internet (ce segment est devenu important en 2005 – année non concernée par l'analyse) et véhicules par région géographique : Nord, Nord-Est, Centre-Ouest, Sud-Est, Sud.

Les noms des véhicules ont été traités pour normalisation et, dans les véhicules homonymes, ont été inclus les lettres de son état (département). De cette manière ont été identifiés "les" Gazeta (AC, ES, RS, SP, MT) ; "Les" Tribuna (AC, SP) ; "Les" Semana (SE, SP) ; "Les" Folha Rural (MG, RS, SP) ; "Les" Jornal da Emater (MG, RS) ; "Les" Jornal do Comercio (PE, RJ) ; "Les" Jornal do Comercio (AM, RS) ; Et "les" Primeira Pagina (AC, SP).

Centres de recherche

Les centres de recherche ont été segmentés par état et par région géographique.

États (départements)

Les états de la fédération brésilienne d'appartenance des véhicules ont été aussi segmentés par région géographique.

Année de publication

Les dates de publication (JJ/MM/AAAA) des nouvelles ont été traitées en sélectionnant d'abord l'année de publication (2003 et 2004) et ensuite, le mois et année de publication (01/2003....12/2004).

Nombre de pages

Cette variable conçue comme "qualitative" - page impaire, page paire, deux pages, trois pages et quatre ou plus pages, été recodée comme "quantitative" - une page, deux pages, trois pages, quatre ou plus pages.

Matrices

La base traitée a été importée dans le logiciel VantagePoint. Les matrices suivantes ont été générées, qui peuvent indiquer des items variés, notés ensuite. Voyez les détails dans le tableau 56.

Tableau 56 : Les matrices générées avec la base Clipping

Totale de nouvelles/centres de recherche (UD) => quels centres de recherche publient le plus.

Totale de nouvelles/état (UF) => quels états publient le plus.

Totale de nouvelles/année de publication (ANO) => l'évolution des nouvelles publiées dans le temps (année).

Totale de nouvelles/date de publication (DTP) => l'évolution des nouvelles publiées dans le temps (mois/année).

Totale de nouvelles/véhicules (VEI) => quels véhicules publient le plus.

Ensuite les croisements de champs :

Mots-clés/états (PCH/UF) => l'évolution des thèmes publiés dans les états.

Véhicules/centres de recherche (VEI/UD) => la politique de publication des journaux ou des centres de recherche.

Véhicules/année de publication (VEI/ANO) => l'évolution par véhicule dans le temps (année ou mois, dans le cas, année).

Véhicules/mots-clés (VEI/PCH) => les thèmes préférés par véhicule.

Véhicules/genre (VEI/GEN) => quel véhicule publie plus d'éditoriaux, articles, reportages, notes,...

Véhicules/présence (VEI/PRE) => quel véhicule dédie quel type d'espace.

Véhicules/nombre de pages (VEI/PAG) => quel véhicule publie plus de pages.

Véhicules/graphiques (VEI/CGR) => quel véhicule donne quel traitement graphique.

Véhicules/source (VEI/FON) => quel véhicule mentionne quelles sources.

Source/présence (FON/PRE) => quelle source a quel type d'espace.

Source/année de publication (FON/ANO) => l'évolution des citations des sources dans le temps (années).

Source/état (FON/UF) => l'évolution des citations des sources par état.

Mots-clés/source (PCH/FON) => quels thèmes préférés par source.

Mots-clés/présence (PCH/PRE) => quels thèmes ont quel type d'espace.

Mots-clés/genre (PCH/GEN) => quels thèmes ont plus éditoriaux, reportages,...

Mots-clés/année de publication (PCH/ANO) => l'évolution des thèmes dans le temps (année).

Genre/année de publication (GEN/ANO) => l'évolution des traitements éditorial dans le temps (année).

Présence/année de publication (PRE/ANO) => l'évolution de la présence éditoriale dans le temps (année).

Graphiques/année de publication (CGR/ANO) => l'évolution des traitements graphique dans le temps (année).

Centres de recherche/genre (UD/GEN) => quels centres de recherche ont quel traitement éditorial.

Centres de recherche/présence (UD/PRE) => quels centres de recherche ont quelle présence éditoriale.

Centres de recherche/graphiques (UD/CGR) => quels centres de recherche ont quel traitement graphique.

Centres de recherche/année de publication (UD/ANO) => l'évolution de l'effort éditorial des centres dans le temps (année).

Il y a plusieurs d'autres matrices qui apportent des informations sur les relations de l'Embrapa avec la presse mais elles n'ont pas été traitées ici par manque d'espace et de temps. Nous avons choisi les matrices les plus illustratives pour ce travail.

Segment véhicules

Les matrices de chaque segment de véhicules sont listées tableau 57.

Tableau 57 : Les matrices des segments de la variable véhicules

Segment véhicules/centres de recherche (VEI/UD) => la politique de publication des journaux du segment ou des centres de recherche envers le segment.
Segment véhicules/année de publication (VEI/ANO) => l'évolution par véhicule du segment dans le temps (année ou mois, dans le cas, année).
Segment véhicules/mots-clés (VEI/PCH) => les thèmes préférés par véhicule du segment.
Segment véhicules/genre (VEI/GEN) => quel véhicule du segment publie plus d'éditoriaux, articles, reportages, nouvelles, notes d'opinion, entre d'autres.
Segment véhicules/présence (VEI/PRE) => quel véhicule du segment dédie quel type d'espace noble.
Segment véhicules/nombre de pages (VEI/PAG) => quel véhicule du segment publie plus de pages.
Segment véhicules/graphiques (VEI/CGR) => quel véhicule du segment donne quel traitement graphique.
Segment véhicules/source (VEI/FON) => quel véhicule du segment mentionne quelles sources.

Résultats

La base Clipping a été évaluée par rapport à la présence de champs vides, ce qui permet d'accéder, selon Quoniam (1996, p.50) à ses particularités et à la qualité des informations présentes dans chaque champ. Quatre références étaient vides. Le tableau 58 fournit les détails.

Tableau 58 : Base Clipping ; Statistique des champs vides

Nom du champ	Intitulé du champ	References non informées	References informées	% vides
Genre	GEN	2209	14790	12,99
Graphiques	CGR	2211	14788	13,00
Presence	PRE	2209	14790	12,97
Source	FON	2232	14337	13,13
Année	ANO	4	16995	0,0002
Date du clipping	DTP	4	16995	0,0002
Nombre de pages	PAG	189	16810	0,01
Mots-clés	PCH	21	16978	0,0001
Cadran de la page	QUA	2803	14196	16,4
Thèmes	TEM	21	16978	0,0001
Titre	TIT	4	16995	0,0002
Centre de recherche	UD	22	16977	0,0001
État	UF	4	16995	0,0002
Vehicule	VEI	4	16995	0,0002

Les 16999 références se sont divisés en 9028 références en 2004 et 7967 en 2003. Des nouvelles publiées en 677 véhicules dans ces deux années ont été analysées. Des 50 véhicules qui ont publié les plus de nouvelles concernant à l'Embrapa et ses centres de recherche, 26 ont augmenté la quantité de matières publiées en 2004. Ils sont marqués avec un astérisque. Parmi eux "O Estado de S. Paulo" (de Sao Paulo, de 366 pour 374), "Jornal de Brasilia" (de Brasilia, de 244 pour 311), "O Popular" (de Goias, de 216 pour 301), "Zero Hora" (de Rio Grande du Sud, de 119 pour 173), "A Tarde" (de Bahia, de 116 pour 128), "O Globo" (de Rio de Janeiro, de 59 pour 82). Voyez les détails dans le tableau 59. Les caractéristiques de chaque véhicule sont signalées en gris. Par exemple, "O Estado de S. Paulo" c'est un journal national et appartient à la région Sud-Est. A remarquer que le premier véhicule de l'agronoçoce "Globo Rural" est le 19^{ème}. L'Embrapa a changé sa couverture de presse depuis 1990. Au début de cette période l'entreprise publiait la majorité de ses nouvelles dans des véhicules spécialisés en agronoçoce. En 2003/2004, entre les 50 premiers véhicules, 45 sont des véhicules d'intérêt général, de grande circulation. Les véhicules de l'agronoçoce sont marqués en gras.

Tableau 59 : Base Clipping ; Les 50 véhicules ayant publié le plus de nouvelles

	2003	2004	Total	Véhicule	Journaux nationaux	Principaux journaux	Journaux des états	Agro- négoce	Nord	Nord- Est	Centre- Ouest	Sud- Est	Sud
1	366	374	740	O Estado de S. Paulo*	1							1	
2	304	293	597	Correio do Povo			1						1
3	244	311	555	Jornal de Brasília*		1					1		
4	310	225	535	Diario da Amazonia			1		1				
5	216	301	517	O Popular*			1				1		
6	308	197	505	Gazeta Mercantil		1						1	
7	152	192	344	A Gazeta-MT*			1				1		
8	183	154	337	Folha do Estado			1				1		
9	64	240	304	Diario da Manha*			1						1
10	159	144	303	Diario do Nordeste			1			1			
11	152	141	293	Meio Norte			1			1			
12	119	173	292	Zero Hora*		1							1
13	153	133	286	Estado de Minas		1						1	
14	162	124	286	O Estadao			1		1				
15	158	97	255	Folha de Rondonia			1		1				
16	140	112	252	Correio Braziliense		1					1		
17	117	128	245	A Tarde*		1				1			
18	124	114	238	Folha de S. Paulo	1							1	
19	125	104	229	Globo Rural				1				1	
20	120	108	228	O Progresso			1				1		
21	94	131	225	DBO*				1				1	
22	120	102	222	O Liberal		1			1				
23	71	131	202	Diario do Comercio e Industria*			1					1	
24	74	119	193	Tribuna do Brasil*			1				1		
25	65	125	190	Folha de Boa Vista*			1		1				
26	88	102	190	Jornal Minuano*			1						1
27	86	96	182	Diario Popular*			1					1	
28	78	100	178	Valor Economico*		1						1	
29	86	81	167	A Critica		1			1				
30	64	102	166	Jornal Diario*			1				1		
31	90	75	165	O Povo			1			1			
32	82	79	161	Cultivar*				1					1
33	84	71	155	O Dia			1			1			
34	64	86	150	Panorama Rural*				1				1	
35	78	63	141	Jornal do Brasil	1							1	
36	59	82	141	O Globo*	1							1	
37	53	82	135	Diario do Povo*			1			1			
38	12	123	135	O Nacional*			1						1
39	31	100	131	Correio do Sul*			1						1
40	65	63	128	Gazeta do Povo			1						1
41	53	70	123	Folha de Londrina*			1						1
42	47	74	121	Primeira Pagina-SP*			1					1	
43	48	72	120	Correio do Estado*			1				1		
44	59	55	114	Jornal do Commercio-RJ			1					1	
45	59	53	112	O Tempo			1					1	
46	52	59	111	Pagina 20*			1		1				
47	40	70	110	A Tribuna-SP*			1					1	
48	65	45	110	Balde Branco				1				1	
49	55	45	100	Correio da Bahia			1			1			
50	55	40	95	A Gazeta-AC			1		1				

Genre journalistique, traitement graphique et type de présence

Quand le volume de nouvelles publiées en 2003 et 2004 est analysé sous l'optique du genre journalistique (type de nouvelle), nous avons remarqué une augmentation en sept des neuf types. Il y a eu chute seulement dans le nombre d'interviews et de lettres du lecteur. Les reportages se sont accru de 330 à 575 (74,2%), les éditoriaux de 77 à 129 (64,9%) et les notes d'opinion de 155 à 238 (53,5%). Les champs vides ont décreu ce qu'indique une amélioration générale de qualité de la base. Les textes sur l'Embrapa représentent à peu près 54% de toutes les publications. Le tableau 60 présente les détails.

Tableau 60 : Base Clipping ; Genre journalistique par année de publication

Total	Type de nouvelle	2003	2004	%
9193	Texte	4324	4869	12,6
2596	Note d'information	1069	1527	42,8
1207	Article	515	692	34,3
905	Reportage	330	575	74,2
393	Note d'opinion	155	238	53,5
224	Lettre du lecteur	118	106	-11,3
206	Éditorial	77	129	64,9
60	Interview	34	26	-30,7
6	Chronique	1	5	500
14790	Total	7967	9028	13,3
2209	ChampVide	1344	861	-56

Note. % = Pourcentage d'accroissement ou de réduction de 2003 par rapport à 2004.

Par rapport au traitement graphique des différents véhicules il y a eu une croissance de 13,3% en tous les items. A remarquer que 57% des publications (des textes, des reportages et des articles pour la plupart) ont un traitement graphique plus distinct ou élaboré ce qui est très

désirable par l'organisation. Les dépêches avec quatre ou plus éléments ont eu un accroissement très important. Les chiffres sont marqués en gras dans le tableau 61.

Tableau 61 : Base Clipping ; Traitement graphique par année de publication

Total	Type de traitement graphique	2003	2004	%
6374	1 élément (texte)	3062	3312	8,1
5132	2 éléments (texte + photo/illustration)	2385	2747	15,1
1335	3 éléments (texte + photo/illustration + box)	519	816	57,2
593	4 éléments (texte + photo + illustration + box)	208	385	85
1354	5 ou plus éléments	451	903	100,2
14788	Total	7967	9028	13,3
2211	ChampVide	1342	865	-55,1

Note. % = Pourcentage d'accroissement ou de réduction de 2003 par rapport à 2004.

Une évolution similaire a été enregistrée pour le type de présence de l'organisation dans les textes journalistiques, il y a eu une croissance de 89,1% des citations en Manchettes de page et de 34,3% en Cape ou Première page. Ils sont marqués en gris au tableau 62.

Tableau 62 : Base Clipping ; Présence éditoriale par année de publication

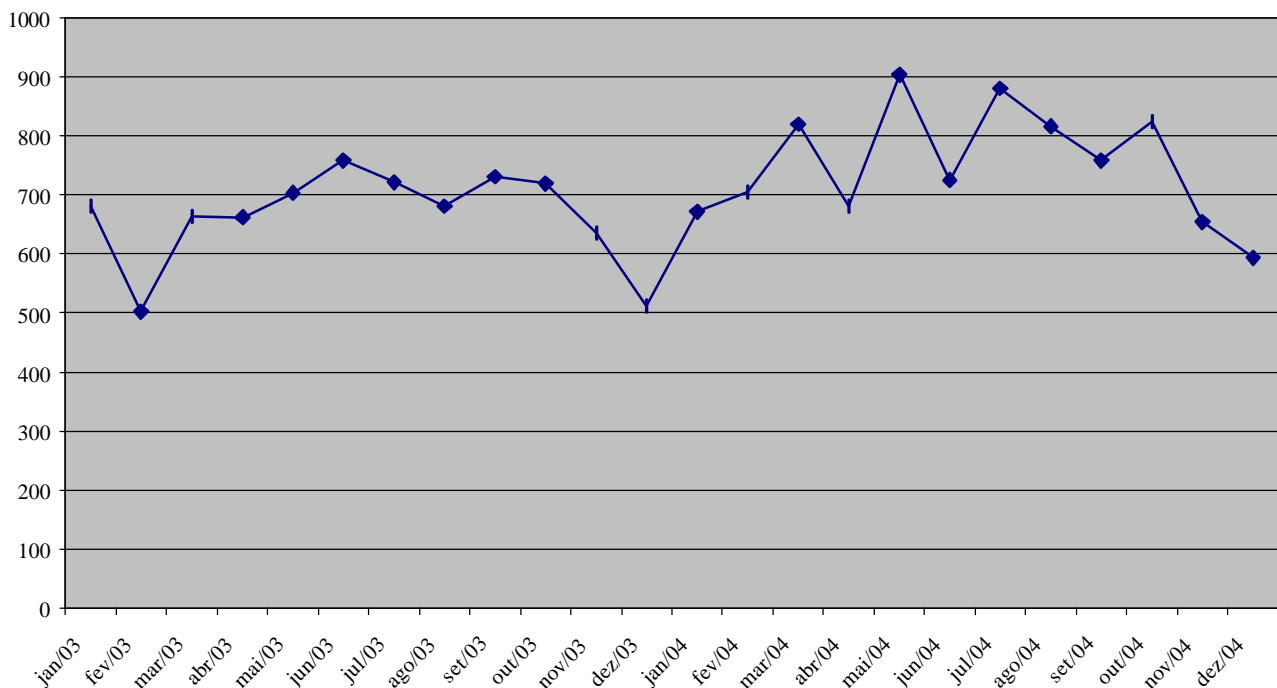
Total	Type de présence	2003	2004	%
9744	Citation	4414	5330	20,7
2053	Titre	904	1149	27,1
1820	Proéminence dans le texte	783	1037	32,4
808	Plinthe/Légende	378	430	13,7
232	Cape/Première Page	99	133	34,3
133	Manchette de page	46	87	89,1
14790	Total	7967	9028	13,3
2209	ChampVide	1343	862	-55,8

Note. % = Pourcentage d'accroissement ou de réduction de 2003 par rapport à 2004.

L'évolution des dépêches pas mois

Par mois, l'évolution générale des dépêches publiées a été plus forte en 2004. La moyenne des publications en 2003 a été de 663,9 nouvelles par mois et en 2004 de 752,3 nouvelles/mois, une augmentation de 6,6%. Cette évolution a fait que pendant toute la période étudiée la moyenne de nouvelles publiées par mois soit de 708,1. Le mois de mai 2004 a enregistré le plus grand nombre de dépêches publiées, 903. Les mois de décembre, janvier et février que coïncident avec les périodes de l'été, des vacances et des fêtes sont ceux les moins productifs en relation au nombre de publications par mois. Le tableau 63 présente les détails.

Tableau 63 : Base Clipping ; Evolution des nouvelles par mois



Les politiques éditoriales des véhicules

L'analyse de la matrice véhicule/nombre d'éléments graphiques présents, relève quatre véhicules qui donnent un traitement graphique très bon aux dépêches sur l'Embrapa : "Jornal de Brasilia" et "O Estado de S. Paulo", deux journaux et "Panorama Rural" et "Globo Rural", deux revues. Méritent aussi d'être cités comme "grands véhicules de l'Embrapa" les revues de l'agronégoce "DBO", "Cultivar", "Balde Branco", "Suinocultura Industrial", "Agroanalysis",

"Rural", "A Granja", "Bahia Agricola", la revue nationale "Veja" et les journaux "O Popular", "O Estado de Minas", "A Tarde", "Folha de S. Paulo", "A Gazeta-MT", "Correio Braziliense", "O Globo", "Diario do Nordeste", "Zero Hora" et "A Critica". A remarquer que quand la variable richesse du traitement graphique est en évidence, les véhicules de l'agronégoce, marqués en gras au tableau 64, sont en force. Ils concèdent le meilleur traitement possible aux informations technico-scientifiques provenant de l'Embrapa. Ces véhicules doivent être rappelés quand l'objectif de l'Embrapa sera de publier des nouvelles avec beaucoup de richesse de détails et/ou d'éléments graphiques (informationnels). Le tableau 64 montre les détails.

Tableau 64 : Véhicules concédant le meilleur traitement graphique à l'Embrapa

	Total	Véhicule	1 élém.	2 élém.	3 élém.	4 élém.	5 ou plus
1	555	Jornal de Brasília	185	167	102	30	71
2	740	O Estado de S. Paulo	298	282	62	28	64
3	150	Panorama Rural	15	51	20	10	54
4	229	Globo Rural	55	109	19	8	36
5	225	DBO	69	78	22	12	36
6	161	Cultivar	22	54	19	13	36
7	517	O Popular	270	162	32	17	31
8	110	Balde Branco	21	39	8	6	31
9	286	Estado de Minas	93	136	20	9	28
10	245	A Tarde	81	98	25	14	25
11	238	Folha de S. Paulo	132	49	22	9	23
12	73	Suinocultura Industrial	13	16	11	2	22
13	344	A Gazeta-MT	161	118	30	8	21
14	252	Correio Braziliense	127	68	23	13	21
15	61	Agroanalysis	11	16	10	4	20
16	60	Revista Rural - SP	12	15	7	6	20
17	141	O Globo	51	52	16	2	19
18	92	A Granja	7	37	17	7	18
19	38	Cultivar Grandes Culturas	2	13	4	1	18
20	27	Bahia Agrícola	4	2	2	2	17
21	33	Cultivar Hortícolas e Frutas	8	6	2	1	16
22	303	Diario do Nordeste	142	81	23	10	15
23	292	Zero Hora	125	48	31	13	15
24	41	Veja	7	4	11	3	15
25	167	A Crítica	86	39	15	9	14
26	121	Primeira Pagina-SP	14	52	19	10	14
27	66	AgriNova	12	22	10	9	13
28	65	Pesquisa Fapesp	8	35	3	6	13
29	45	Produtor Rural	9	18	4	1	13
30	22	Epoca	2	4	2	1	13
31	63	Produtor Parmalat	11	28	5	3	12
32	59	Avicultura Industrial	20	15	8	1	12
33	53	Anuario Brasileiro da Uva e do Vinho	1	18	11	11	12
34	337	Folha do Estado	77	211	29	9	11
35	293	Meio Norte	174	59	20	5	11
36	92	Hoje em Dia	41	29	8	3	10
37	63	Safra	10	21	7	3	10
38	24	Revista Item	1	10	2		10
39	17	Anuario Brasileiro do Arroz			1	2	10
40	23	IstoE	2	5	7		9

Note. élém. = éléments graphiques présents dans le texte.

Quand nous analysons les véhicules qui publient des dépêches avec seulement un élément (texte) affleurent des considérations sur leur politique éditoriale. Les deux journaux qui apparaissent dans les deux premières places, "Correio do Povo" et "Gazeta Mercantil" sont aussi des journaux qui ont comme politique éditoriale, publier seulement des textes. Les deux doivent être considérés comme "grands véhicules de l'Embrapa" ainsi que le "Diario da Amazonia", "Meio Norte", "O Liberal", "A Gazeta-MT" et suivantes. Le tableau 65 exhibe les détails.

Tableau 65 : Véhicules publiant le plus de dépêches simples sur l'Embrapa

	Total	Véhicule	1 élém.	2 élém.	3 élém.	4 élém.	5 ou plus
1	597	Correio do Povo	399	164	9	2	3
2	505	Gazeta Mercantil	345	102	25	9	7
3	740	O Estado de S. Paulo	298	282	62	28	64
4	517	O Popular	270	162	32	17	31
5	535	Diario da Amazonia	252	204	26	5	5
6	555	Jornal de Brasilia	185	167	102	30	71
7	293	Meio Norte	174	59	20	5	11
8	222	O Liberal	164	31	12	7	6
9	344	A Gazeta-MT	161	118	30	8	21
10	303	Diario do Nordeste	142	81	23	10	15
11	193	Tribuna do Brasil	139	38	9	3	4
12	238	Folha de S. Paulo	132	49	22	9	23
13	252	Correio Braziliense	127	68	23	13	21
14	292	Zero Hora	125	48	31	13	15
15	202	Diario do Comercio e Industria	120	55	24	2	
16	286	O Estadao	116	140	4	2	3
17	190	Folha de Boa Vista	115	64	5	3	2
18	178	Valor Economico	113	39	14	5	6
19	286	Estado de Minas	93	136	20	9	28
20	141	Jornal do Brasil	89	37	7	3	5
21	167	A Critica	86	39	15	9	14
22	304	Diario da Manha	83	49	7	5	4
23	245	A Tarde	81	98	25	14	25
24	182	Diario Popular	81	40	4		2
25	135	Diario do Povo	81	24			
26	114	Jornal do Commercio-RJ	78	22	7	3	3
27	337	Folha do Estado	77	211	29	9	11
28	155	O Dia	76	44	8	7	6
29	165	O Povo	75	53	9	4	5
30	255	Folha de Rondonia	70	88	1		
31	225	DBO	69	78	22	12	36
32	128	Gazeta do Povo	65	27	16	8	8
33	229	Globo Rural	55	109	19	8	36
34	112	O Tempo	52	40	14	2	4
35	190	Jornal Minuano	51	24	20	5	1
36	141	O Globo	51	52	16	2	19
37	100	Correio da Bahia	50	40	6	2	2
38	82	Diario Catarinense	48	23	4	3	3
39	123	Folha de Londrina	41	34	9		
40	92	Hoje em Dia	41	29	8	3	10

Note. élém. = éléments graphiques présents dans le texte.

Les véhicules qui mentionnent le plus l'entreprise en leurs premières pages sont : "Jornal de Brasilia" (DF), "O Estado de S. Paulo" (SP), "O Popular" (GO), "Tribuna do Brasil" (DF), "Diario da Amazonia" (RO), "Gazeta Mercantil" (SP), "Primeira Pagina-SP" et "A Tarde". Plusieurs de ces journaux sont aussi parmi ceux qui donnent le meilleur traitement graphique (informationnel ou nombre de textes) aux dépêches de l'Embrapa. Ils sont marqués en gras au tableau 66.

Le "Folha de S. Paulo" (SP) a été le véhicule qui a le plus cité l'Embrapa dans leurs manchettes. Il est marqué en gris comme d'autres journaux avec la même caractéristique, jusqu'à la fréquence cinq comme montré dans le tableau 66.

Tableau 66 : Véhicules ayant cité le plus à l'Embrapa dans leur première page

	Total	Véhicule	Première Page	Manchette
1	555	Jornal de Brasilia	21	5
2	740	O Estado de S. Paulo	11	5
3	517	O Popular	10	
4	193	Tribuna do Brasil	9	1
5	535	Diario da Amazonia	8	7
6	505	Gazeta Mercantil	8	2
7	121	Primeira Pagina-SP	8	
8	245	A Tarde	8	
9	178	Valor Economico	7	3
10	110	A Tribuna-SP	6	
11	286	Estado de Minas	5	3
12	190	Folha de Boa Vista	5	1
13	238	Folha de S. Paulo	4	10
14	286	O Estadao	4	8
15	141	Jornal do Brasil	4	2
16	202	Diario do Comercio e Industria	4	1
17	597	Correio do Povo	4	1
18	12	Jornal Ciencia e Tecnologia	4	

L'Éditorial est la place où un véhicule exprime son opinion. C'est pourtant un espace noble. Les journaux et revues qui ont plus cité l'Embrapa dans leurs éditoriaux sont : "Diario Popular", "O Estado de S. Paulo", "Gazeta Mercantil", "Jornal do Brasil", "Panorama Rural-SP", "A Gazeta-MT", "Gazeta do Povo" et "O Globo". Le tableau 67 présente les détails.

Tableau 67 : Véhicules publiant le plus d'éditoriaux sur l'Embrapa

	Total	Véhicule	Editorial	Texte
1	182	Diario Popular	13	72
2	740	O Estado de S. Paulo	10	323
3	505	Gazeta Mercantil	10	345
4	141	Jornal do Brasil	9	68
5	150	Panorama Rural	7	63
6	344	A Gazeta-MT	6	218
7	141	O Globo	5	74
8	128	Gazeta do Povo	5	92
9	303	Diario do Nordeste	4	175
10	293	Meio Norte	4	198
11	66	Agrinova	4	38
12	73	O Estado do Maranhao	4	56
13	252	Correio Braziliense	4	101

Quand nous regardons les deux autres genres journalistiques avec le plus grand taux de croissance, la classification change. Les véhicules publiant le plus de Reportages sont, en leur majorité, les spécialisés en Agronégoce. Ils sont marqués en gris clair dans le tableau 68. Pour les Notes d'opinion, les véhicules de l'agronégoce laissent la place aux quotidiens d'intérêt général. Plusieurs de ces quotidiens sont présents dans les deux genres, ils publient aussi beaucoup de reportages sur l'Embrapa. Ils sont marqués en gras dans le tableau 68. A remarquer enfin que la principale revue hebdomadaire du Brésil, "Veja" (marquée en gris plus fort dans le tableau 68) est classée entre les véhicules qui publient le plus de reportages.

Tableau 68 : Véhicules publiant le plus de reportages et notes d'opinion sur l'Embrapa

	Total	Véhicule	Reportage	Total	Véhicule	Note d'Opinion
1	150	Panorama Rural	39	555	Jornal de Brasilia	40
2	740	O Estado de S. Paulo	37	252	Correio Braziliense	31
3	229	Globo Rural	29	141	Jornal do Brasil	23
4	555	Jornal de Brasilia	22	141	O Globo	20
5	286	Estado de Minas	22	535	Diario da Amazonia	16
6	110	Balde Branco	22	303	Diario do Nordeste	15
7	121	Primeira Pagina-SP	19	222	O Liberal	15
8	20	Anuario Brasileiro da Pecuaria	19	740	O Estado de S. Paulo	14
9	167	A Critica	18	238	Folha de S. Paulo	12
10	65	Correio Popular	17	165	O Povo	12
11	19	Anuario Brasileiro Aves e Suinos 2003	17	92	Hoje em Dia	12
12	517	O Popular	16	193	Tribuna do Brasil	11
13	123	Folha de Londrina	16	293	Meio Norte	10
14	92	A Granja	16	167	A Critica	10
15	60	Revista Rural - SP	16	63	Correio da Paraiba	10
16	57	Jornal do Comercio-AM	15	155	O Dia	7
17	225	DBO	14	292	Zero Hora	7
18	141	O Globo	13	73	O Estado do Maranhao	6
19	238	Folha de S. Paulo	13	505	Gazeta Mercantil	5
20	41	Veja	13	517	O Popular	5
21	303	Diario do Nordeste	12	112	O Tempo	5
22	182	Diario Popular	12	38	Cultivar Grandes Culturas	5
23	292	Zero Hora	11	51	Gazeta de Alagoas	4
24	505	Gazeta Mercantil	11	51	Jornal do Comercio-PE	4
25	45	Produtor Rural	11	57	Jornal do Comercio-AM	4

Les polarisations des éditoriaux de deux journaux nationaux, une approximation

Comme la variable polarisation (positive ou négative) n'a pas été informée dans cette base Clipping nous avons examiné les titres et les contenus des éditoriaux de deux journaux nationaux, "O Estado de S. Paulo" et "O Globo". Les dix éditoriaux de "O Estado de S. Paulo" étaient titrés : *"Emergencia na Embrapa [Emergence à l'Embrapa] ; Ma-fe e ineptia [Mauvaise-foi et ineptie] ; Ciencia pouco aplicada [Science peu appliquée]; A reforma que Dirceu promete [Le changement que Dirceu promet] ; E preciso preservar a Embrapa [Il faut préserver l'Embrapa] ; Um pedaco brasileiro de Primeiro Mundo [Un morceau brésilien de*

Premier Monde] ; *Boa safra e boas exportacoes* [Bonne récolte et bonnes exportations] ; *Em defesa dos transgenicos* [Pour les transgéniques] ; *A hora da inovacao, segundo Palocci* [L'heure de l'innovation, d'après Palocci] ; *A 'Embrapa industrial'* [L'Embrapa industrielle] ". Des dix, seulement un, "*Emergence à l'Embrapa*" avait une connotation défavorable ayant relation avec une critique du journal aux politiques de la direction antérieure de l'entreprise. Dans l'"O Globo" : "*Cerco ideologico* [Encerclement idéologique] ; *Na contramao* [À contre-courant] ; *Ciencia livre* [Science libre] ; *Ciencia a salvo* [Science sauvée] et *Dogmas* [Des dogmes]" tous sont favorables à l'Embrapa.

Les caractéristiques des couvertures

Quand la variable "centre de recherche" de l'Embrapa est segmentée par états et croisée avec la variable véhicules, aussi segmentée par états, les résultats indiquent que la majorité des journaux tend à publier plus de dépêches des centres de recherche de leurs états. Par exemple, les journaux du Rio Grande do Sul (RS) publient plus de nouvelles des centres de recherche du Rio Grande du Sud ; Les journaux de l'Amazonas (AM) tendent à publier plus de nouvelles du centre de l'Amazonas.

Ces résultats n'expriment pas seulement les politiques éditoriales des véhicules mais sont aussi très fidèles aux efforts de divulgation et communication de l'Embrapa qui misent sur des efforts de communication nationaux au siège et régionales, étatiques et locales aux centres de recherche.

Les états marqués avec des colonnes en gris plus clair dans le tableau 69, dans ; Mato Grosso (MT), Tocantins (TO), Maranhão (MA), Rio Grande do Norte (RN), Alagoas (AL) et Espírito Santo (ES) se caractérisent par le fait qu'il n'y a pas des centres de recherche de l'Embrapa. Les cellules marquées en gris dans le tableau 69 sont celles des croisements état-état. Si nous ne considérons pas la couverture du Siège qui est faite avec la participation de tous les centres de recherche pour des sujets "nationaux", en cinq états qui ont leurs abréviations marquées aussi en gris dans le tableau 69, Rio Grande do Sul (RS), Parana (PR), Piauí (PI), Acre (AC) et Sergipe (SE) le nombre de dépêches des centres est plus grand que les nouvelles du Siège de l'Embrapa. En seulement deux de ces croisements, Goiás (GO) et Rio de Janeiro (RJ), les effectifs les plus grands ne sont pas aux croisements états-états. La proximité de l'état de Goiás et du District Fédéral (DF) et de leurs conditions écologiques bien comme le fait qu'il y ait six unités de l'Embrapa au District Fédéral contre une seule au Goiás explique en partie la

prédominance des nouvelles du DF dans ce cas. Par contre, l'état du Rio de Janeiro (RJ) nous semble mériter une analyse plus complète. Pour cela nous avons marqué la colonne et la ligne Rio de Janeiro avec un gris plus fort dans le tableau 69.

Il y a trois centres de recherche de l'Embrapa au Rio de Janeiro, l'Embrapa Sols (CNPS), l'Embrapa Agro-industrie d'Aliments (CTAA) et l'Embrapa Agrobiologie (CNPAB) et l'état est aussi le siège de deux importants journaux nationaux brésiliens, "O Globo" et "Jornal do Brasil". Malgré cela, les trois états qui publient le plus de dépêches des centres du Rio sont l'état de Sao Paulo (SP) et le District Fédéral (DF) avec le Rio Grande do Sul (RS) et le croisement Rio x Rio en troisième place. Ils sont à la ligne horizontale dans le tableau 69.

Tableau 69 : Centres de recherche de l'Embrapa par état et véhicules par état

Total	Centres Véhicules	SP	RS	DF	RO	MT	MG	GO	MS	PR	PI	CE	RJ	BA	AM	PA	SC	AC	RR	SE	TO	PB	PE	MA	RN	AL	ES	AP
		7990	Siège	2384	875	601	488	391	372	239	132	248	231	317	316	212	218	180	112	95	134	50	82	57	67	55	45	43
1516	RS	256	1098	28	11	16	12	24	3	16	3		13	6		1	25						1			1	1	
1267	DF	291	113	368	46	72	68	109	6	16	11	14	41	28	17	2	5	6	8	9	17	3	6	2	4	2	1	2
1110	SP	886	33	46	10	28	13	22	6	8	1	4	8	8	7	1	6	7	2	6	1	1		4	3	2		
1004	PR	285	181	35	34	47	39	34	7	274	1	2	10	18		1	3	1	2	2	9	4		7	2	5	1	
914	MS	200	59	33	41	56	23	37	433	4		1	7	7	1	2		4		1	2			1	1	1		
648	MG	331	33	23	23	25	131	44	5	11	1	1	2	2	2	1	1	1		4	1		5		1			
470	RO	15	8	3	418	9	3	2	1	1				1				5	2				1	1				
389	SC	214	31	3	5	5	3	4	1	8			7	1			102	2					1	1				1
361	CE	86	15	16	4	7	8	4		2	28	161	1	5	1		2	1	1				3	1	3	11	1	
357	PI	29	31	3	1	1	1		2		271	2	1	2		1		1		1			1	9				
251	GO	64	22	26	11	10	10	83	2	1	3		6	2				3		1	6		1					
235	BA	52	18	13	12	14	10	13	1		3	2	7	74					1	1	5	2	1	5	1			
219	AC	26	2	8	43	12	4	3	1	1		1	1	1	4			110						1				1
226	RJ	73	25	31	15	6	8	7	7	4	2	2	25	2	5	2		2	1	2			4	1	2			
195	PA	44	6	8	22	10				1			1		6	73	1	12	5	5			1					
185	SE	50	3	8	3	2	2	4		1	3		1	7	1					96			1	1		2		
183	PB	26	18	6	4	26	10	22	2	3	2	9	3	8	1	1		1	1	1	2	35	1		1			
151	PE	31	20	12	10	5	3	3		1	6	5	4	29	1		1			3			15	1	1			
132	AM	20	2	4	23	5	3	1						3	67	1		1	1					1				
88	RR	9	4	3	22	16		1		1								2	28						1		1	
15	AP	2	1	1	7	1																						3
17906	Total	5376	2598	1279	1253	764	723	656	609	601	566	521	454	416	331	266	258	254	186	182	125	106	107	92	73	56	43	15

Une étude du Rio de Janeiro

Nous avons voulu étudier de plus près le phénomène Rio de Janeiro. Nous avons cherché les véhicules de l'état du Rio de Janeiro qui avaient publié, dans les deux ans, les 25 dépêches des centres Embrapa du Rio de Janeiro. Ils sont onze. Les résultats sont fournis tableau 70.

Tableau 70 : Véhicules de l'état du Rio de Janeiro ayant publié des dépêches des centres de recherche de l'Embrapa du Rio

Véhicules du Rio	Total
O Globo	6
Jornal do Comercio-RJ	4
Jornal do Brasil	3
A Lavoura	3
Ciencia Hoje	2
Monitor Mercantil	2
Diario de Macae	1
Gazeta de Itaperuna e Regiao	1
Rio Inteligente	1
Rural Semanal - RJ	1
Valor Economico	1
Total	25

A remarquer qu'il y a des véhicules du Rio de Janeiro comme "Extra", "O Dia" (les deux journaux qui vendent le plus dans cet état), "Agroanalysis" et "O Fluminense" qui n'ont publié aucune nouvelle des centres de recherche de Rio. Ces informations ont été comparées avec les véhicules qui publient des dépêches des centres de recherche du Rio de Janeiro. Là une petite surprise. Le journal "Direto do Cerrado" (14, District Fédéral, DF) est celui que publie le plus de nouvelles de ces centres. Mais il y a dans cette liste "O Estado de S. Paulo" (11, Sao Paulo, SP), "A Granja" (10, Rio Grande do Sul, RS), "Diario da Amazonia" (9, Rondônia, RO) et "Jornal de Brasilia" (8, District Fédéral, DF) avant d'arriver au premier véhicule du Rio de Janeiro.

Plusieurs explications sont plausibles pour ce phénomène : D'abord, internes : La couverture déficiente de la base Clipping par rapport aux véhicules du Rio de Janeiro, la difficulté des thèmes traités par les centres en termes de publication ou un travail insatisfaisant de communication avec les journaux du Rio. Ensuite, externes : L'importance comparative de l'agronégoce au Rio de Janeiro. Les deux états concernés, Sao Paulo et Rio Grande do Sul et le District Fédéral sont des places importantes de l'agronégoce brésilien. Alors, leurs journaux

dédient plus d'attention aux sujets de l'agriculture, de l'élevage, des forêts et des ressources du sol que ceux du Rio de Janeiro.

De toute façon, ces résultats sont mis en évidence pour des nécessaires ajustements dans les politiques de divulgation des centres de recherche du Rio de Janeiro. Le tableau 71 présente la liste des 26 véhicules qui ont publié jusqu'à trois nouvelles des centres du Rio en 2003 et 2004. Les véhicules du Rio sont marqués en gras.

Tableau 71 : Les 26 véhicules ayant publié jusqu'à trois nouvelles des centres du Rio

Véhicule	Total
1 Direto do Cerrado	14
2 O Estado de S. Paulo	11
3 A Granja	10
4 Diario da Amazonia	9
5 Jornal de Brasilia	8
6 O Globo	6
7 Revista Item	5
8 O Estadao	5
9 Diario do Comercio e Industria	5
10 O Popular	5
11 Cultivar	4
12 Tribuna do Brasil	4
13 Gazeta Mercantil	4
14 Safra	4
15 Jornal do Commercio-RJ	4
16 Correio do Povo	4
17 Revista Rural - SP	3
18 Jornal do Brasil	3
19 Voz do Parana	3
20 Estado de Minas	3
21 A Lavoura	3
22 A Gazeta-MT	3
23 Pesquisa Fapesp	3
24 Folha do Estado	3
25 Panorama Rural	3
26 Globo Rural	3
Total 26 premiers	132
Total RJ	226
% 26 premiers	60,5

Les politiques éditoriales par segment de véhicules

Avertis de la question "carioca" (ceux qui naissent à Rio) nous tournons notre attention pour les caractéristiques des politiques éditoriales des journaux nationaux et des principaux journaux, dont l'influence est régionale. Pour cela nous avons d'abord segmenté les centres de recherche par région et les véhicules par journaux nationaux. Ces journaux sont vendus dans tout le territoire brésilien et devraient par définition avoir une vision plus "nationale" et moins "régionale" ou même "étatique".

Dans l'analyse du tableau 69 nous avons constaté, dans la couverture de l'Embrapa, une tendance géographique de publication la presse, de publier plus de dépêches des centres de recherche les plus proches. Cette tendance apparaît encore dans le tableau 72. Les journaux nationaux de Sao Paulo (SP) affichent une couverture plus éclectique du territoire brésilien, en contrastant avec celle des journaux du Rio. "O Globo" et "Jornal do Brasil" ont publié une nouvelle chacun sur les centres du Nord-Est, aucune des centres du Nord et cinq nouvelles des centres de la région Sud, sept au total. "O Estado de S. Paulo" et "Folha de S. Paulo" ont publié 41 nouvelles des centres du Nord-Est, 16 des centres du Nord et 100 des centres de l'Embrapa de la région Sud, un total de 157 nouvelles. Le tableau 72 exhibe les détails.

Tableau 72 : Centres par région et les journaux nationaux

Total par région	Centres par Région \ Journaux nationaux	O Estado de S. Paulo	Folha de S. Paulo	O Globo	Jornal do Brasil	Total par journal
7990	Siège	355	166	126	119	766
2847	Sud	76	24	3	2	105
2358	Centre-Ouest	96	28	5	17	146
1944	Sud-Est	196	15	7	3	221
1405	Nord-Est	32	9	1	1	43
1035	Nord	13	3			16
17579		740	238	141	141	1260

L'analyse des politiques éditoriales des principaux journaux nous apporte une comparaison intéressante dans le tableau 73, entre les deux principaux journaux de Brasilia (DF), "Correio Braziliense" et "Jornal de Brasilia". La couverture du deuxième est plus nationale et

équilibrée. Le premier a publié six nouvelles de la région Sud, cinq du Sud-Est, deux du Nord-Est et une du Nord, 14 au total. Le deuxième a publié respectivement 38, 44, 30 et 12 nouvelles, 124 au total. En comparant les deux journaux économiques, "Gazeta Mercantil" et "Valor Economico", nous vérifions l'absence de dépêches des centres de recherche du Nord et du Nord-Est dans le **Valor** pendant que le **Gazeta** a publié respectivement 48 et 13 dépêches des centres de recherche de ces régions, 61 au total. Le "Zero Hora" principal journal du Rio Grande do Sul (RS) a publié 3 nouvelles en 2003 et 2004 concernant les centres de l'Embrapa des régions Nord et Nord-Est. En compensation, "O Liberal" et "A Critica" les principaux journaux du Nord, respectivement des états du Para (PA) et de l'Amazonas (AM) n'ont publié aucune nouvelle provenant des centres de l'Embrapa de la région Sud et trois nouvelles sur les centres du Nord-Est. Toutes ces cellules sont marquées en gris dans le tableau 73.

Tableau 73 : Centres par région et les principaux journaux

Total par région	Centres par région \ Principaux Journaux	Jornal de Brasilia	Correio Braziliense	Estado de Minas	Gazeta Mercantil	Valor Economico	A Tarde	Zero Hora	O Liberal	A Critica
7990	Siège	180	197	155	342	144	115	170	155	110
2847	Sud	38	6	24	26	13	18	109		
2358	Centre-Ouest	292	46	43	56	14	30	11	5	6
1944	Sud-Est	44	5	53	25	6	8	5	4	8
1405	Nord-Est	30	2	15	48		79	1	1	2
1035	Nord	12	1	3	13		4	2	60	46
17579	Total	555	252	286	505	178	245	292	222	167

Les efforts éditoriaux des centres de recherche

Quantifiant l'effort éditorial des centres de recherche et du Siège dans les deux années étudiées nous constatons que douze centres présentent une réduction des quantités de dépêches publiées. Ils sont marqués en gris. Vingt-huit centres ont augmenté leur divulgation de 2003 à 2004. Le tableau 74 présente les détails. A remarquer que les centres de recherche du Rio de Janeiro sont dans la partie inférieure du tableau.

Tableau 74 : Centres par nombre de nouvelles et année

	Total	Centre de recherche	2003	2004	03_04
1	7990	Siège Embrapa	3805	4185	+
2	901	Embrapa Soja (PR)	352	549	+
3	624	Embrapa Trigo (RS)	154	470	+
4	604	Embrapa Pecuaria Sudeste (SP)	258	346	+
5	542	Embrapa Agropecuaria Oeste (MS)	265	277	+
6	470	Embrapa Rondonia (RO)	292	178	
7	462	Embrapa Cerrados (DF)	220	242	+
8	455	Embrapa Gado de Leite (MG)	280	175	
9	443	Embrapa Recursos Geneticos e Biotecnologia (DF)	218	225	+
10	417	Embrapa Clima Temperado (RS)	156	261	+
11	389	Embrapa Suinos e Aves (SC)	186	203	+
12	357	Embrapa Meio-Norte (PI)	170	187	+
13	325	Embrapa Gado de Corte (MS)	187	138	
14	264	Embrapa Uva e Vinho (RS)	117	147	+
15	254	Embrapa Pecuaria Sul (RS)	109	145	+
16	251	Embrapa Arroz e Feijao (GO)	100	151	+
17	235	Embrapa Mandioca e Fruticultura (BA)	140	95	
18	230	Embrapa Agroindustria Tropical (CE)	136	94	
19	219	Embrapa Acre (AC)	114	105	
20	212	Embrapa Hortalias (DF)	94	118	+
21	199	Embrapa Milho e Sorgo (MS)	98	101	+
22	197	Embrapa Monitoramento por Satelite (SP)	102	95	
23	195	Embrapa Amazonia Oriental (PA)	93	102	+
24	185	Embrapa Tabuleiros Costeiros (SE)	81	104	+
25	183	Embrapa Algodao (PB)	94	89	
26	156	Embrapa Meio Ambiente (SP)	84	72	
27	151	Embrapa Semi-Arido (PE)	69	82	+
28	140	Embrapa Instrumentacao Agropecuaria (SP)	57	83	+
29	138	Embrapa Caprinos (CE)	75	63	
30	132	Embrapa Amazonia Ocidental (AM)	66	66	
31	129	Embrapa Solos (RJ)	58	71	+
32	108	Embrapa Florestas (PR)	69	39	
33	102	Embrapa Informacao Tecnologica (DF)	45	57	+
34	88	Embrapa Roraima (RR)	29	59	+
35	82	Embrapa Transferencia de Tecnologia (DF)	35	47	+
36	71	Embrapa Pantanal (MS)	39	32	
37	68	Embrapa Cafe (DF)	25	44	+
38	60	Embrapa Agrobiologia (RJ)	17	43	+
39	54	Embrapa Informatica Agropecuaria (SP)	20	34	+
40	37	Embrapa Agroindustria de Alimentos (RJ)	14	23	+
41	18	ChampVide	13	5	
42	15	Embrapa Amapa (AP)	7	8	+

Embrapa Soja, CNPSO (901), Embrapa Blé, CNPT (624) et Embrapa Élevage du Bétail Sud-Est, CPPSE (604) sont les centres de recherche qui ont publié le plus de dépêches dans la presse pendant la période étudiée.

Le tableau suivant, des centres de recherche par les véhicules montre deux choses. La première, quand la matrice est ordonnée par le nombre de nouvelles publiées par centre, nous permet d'identifier les véhicules qui publient plus de nouvelles d'un centre spécifique. Les dépêches du Siège de l'entreprise, par exemple, sont publiés en préférence au "O Estado de S. Paulo" (355), "Gazeta Mercantil" (342), "Correio do Povo" (290), "Diario da Amazonia" (232) et "O Popular" (203). La liste de véhicules de l'Embrapa Soja est différente; "Folha de Londrina" (82), "Jornal de Londrina" (58), "Cultivar" (36), "O Estado de S. Paulo" (30) et "Correio do Povo" (24). Chaque centre de recherche a sa liste de véhicules préférentiels. Le tableau 75 montre les détails.

Tableau 75 : Véhicules par Centre ; Siège Embrapa (DF), Embrapa Soja (PR), Embrapa Blé (RS), Embrapa Bovins de Lait (MG), Embrapa Élevage du Bétail Sud-Est (SP) et Embrapa Sols (RJ), vision 1

Total	Véhicule	Siège Embrapa (DF)	Total	Véhicule	Embrapa Soja (PR)	Total	Véhicule	Embrapa Blé (RS)
740	O Estado de S. Paulo (SP)	355	123	Folha de Londrina (PR)	82	304	Diario da Manha (RS)	162
505	Gazeta Mercantil (SP)	342	69	Jornal de Londrina (PR)	58	135	O Nacional (RS)	118
597	Correio do Povo (RS)	290	161	Cultivar (RS)	40	597	Correio do Povo (RS)	76
535	Diario da Amazonia (RO)	232	740	O Estado de S. Paulo (SP)	30	292	Zero Hora (RS)	42
517	O Popular (GO)	203	597	Correio do Povo (RS)	25	161	Cultivar (RS)	18
252	Correio Braziliense (DF)	197	150	Panorama Rural (SP)	25	740	O Estado de S. Paulo (SP)	12
303	Diario do Nordeste (CE)	194	128	Gazeta do Povo (PR)	24	555	Jornal de Brasilia (DF)	9
344	A Gazeta-MT (MT)	182	337	Folha do Estado (MT)	23	517	O Popular (GO)	9
555	Jornal de Brasilia (DF)	180	39	Anuario Brasileiro da Soja (RS)	22	13	Folha Rural-RS (RS)	8
337	Folha do Estado (MT)	176	36	Informativo Meridional (PR)	22	505	Gazeta Mercantil (SP)	7

Total	Véhicules	Embrapa Bovins de Lait (MG)	Total	Véhicules	Embrapa Élevage du Bétail Sud-Est (SP)	Total	Véhicules	Embrapa Sols (RJ)
225	DBO (SP)	62	92	A Folha (SP)	72	53	Direto do Cerrado (DF)	13
110	Balde Branco (SP)	48	121	Primeira Pagina-SP (SP)	62	92	A Granja (RS)	8
517	O Popular (GO)	22	110	A Tribuna-SP (SP)	59	740	O Estado de S. Paulo (SP)	8
286	Estado de Minas (MG)	17	225	DBO (SP)	53	535	Diario da Amazonia (RO)	6
19	Leite DPA (SP)	16	110	Balde Branco (SP)	53	24	Revista Item (SP)	5
740	O Estado de S. Paulo (SP)	15	740	O Estado de S. Paulo (SP)	35	161	Cultivar (RS)	3
63	Produtor Parmalat (SP)	14	229	Globo Rural (SP)	24	60	Revista Rural - SP (SP)	3
12	Leite Nestle (SP)	12	555	Jornal de Brasilia (DF)	20	193	Tribuna do Brasil (DF)	3
535	Diario da Amazonia (RO)	11	63	Produtor Parmalat (SP)	19	202	Diario do Comercio e Industria (SP)	3
20	Anuario Brasileiro da Pecuaria (RS)	11	27	Sao Carlos News (SP)	18	141	Jornal do Brasil (RJ)	3

Tous ces tableaux peuvent être utilisés par un centre de recherche pour créer une liste d'importance de véhicules, pour définir priorités pour une divulgation et pour la suggestion de reportages spéciaux. Le chargé de la communication au centre de recherche peut identifier rapidement ces "véhicules préférentiels", voir l'évolution de la couverture en chaque véhicule dans le temps et procéder à des corrections de cours ou d'emphase de couverture. Dans ce cas, il faut travailler pour combler les espaces vides. Le tableau 76 présente l'exemple de la liste des véhicules de l'Embrapa Soja.

Tableau 76 : Évolution de la couverture par véhicule par centre ; Ex. Embrapa Soja

Total	Véhicule	déc/04	nov/04	oct/04	sep/04	aoû/04	jul/04	jun/04	mai/04	avr/04	mar/04	fév/04	jan/04	
1	82	Folha de Londrina	6	5	6	3	2	8	4	2	4	3	4	2
2	58	Jornal de Londrina	3	3	7	2	2	5	2	5	2	3	2	
3	25	Correio do Povo	2	2	3	1				1	2	1	1	1
4	24	Gazeta do Povo	4	1	2	1				3	1	4		
5	23	Folha do Estado			3		1	1	1	1	2	1		3
6	21	A Gazeta-MT			1	1	1	2	1	2	1		1	2
7	21	O Popular		2			2	1		1	2	3	4	
8	19	Mercosul								2	12	3		
9	14	Diario da Amazonia	1		2		1			1	1	3		
10	12	Diario do Comercio e Industria	3	1			1	2	1	1				1
11	11	Folha de Rondonia	1		1			2						1
12	10	Tribuna do Brasil			2		1	1		1	1	2		
13	6	Diario da Manha					1	5						
14	6	Hoje em Dia										1	2	
15	5	O Estado do Maranhao	1											
16	5	Gazeta de Alagoas			1		1	1						
17	4	O Nacional						2						
18	3	Correio da Bahia									1			1
19	3	O Diario												
20	3	O Estadao						1						1

La deuxième chose, quand la matrice est ordonnée par le nombre de nouvelles publiées par journal, montre les centres de recherche qui publient le plus en chaque journal. "O Estado de S. Paulo", par exemple, publie pratiquement la moitié de la production de l'Embrapa Surveillance par Satellite (**Embrapa Monitoramento por Satellite, SP**). Autres centres qui apparaissent sont l'Embrapa Élevage du Bétail Sud-Est (**Embrapa Pecuaria Sudeste, SP**), Embrapa Soja (PR), Embrapa Ressources Génétiques et Biotechnologie (**Embrapa Recursos Geneticos e Biotecnologia, DF**) et Embrapa Bovins de Viande (**Embrapa Gado de Corte, MS**). Le tableau 77 présente les détails.

Tableau 77 : Centres par véhicule, "O Estado de S. Paulo", vision 2

Total	Centre de recherche \ Véhicule	O Estado de S. Paulo
7990	Embrapa Sede	355
197	Monitoramento por Satelite	104
604	Pecuaria Sudeste	35
901	Soja	30
443	Recursos Geneticos e Biotecnologia	29
325	Gado de Corte	19
389	Suinos e Aves	16
199	Milho e Sorgo	16
455	Gado de Leite	15
462	Cerrados	14
156	Meio Ambiente	13
624	Trigo	12
212	Hortalicas	11
417	Clima Temperado	11
235	Mandioca e Fruticultura	10

De l'autre côté le "Gazeta Mercantil" publie plus de dépêches de l'Embrapa Agro-industrie Tropicale (**Embrapa Agroindustria Tropical**, CE), Embrapa Soja (PR), Embrapa Cerrados (CPAC, DF), Embrapa Ressources Génétiques et Biotechnologie (**Embrapa Recursos Geneticos e Biotecnologia**, DF) et Embrapa Riz et Haricot (**Embrapa Arroz e Feijao**, GO). Le tableau 78 présente ces chiffres.

Tableau 78 : Centres par véhicule, "Gazeta Mercantil", vision 2

Total	Centre de recherche \ Véhicule	Gazeta Mercantil
7990	Embrapa Sede	342
230	Agroindustria Tropical	27
901	Soja	14
462	Cerrados	14
443	Recursos Geneticos e Biotecnologia	11
251	Arroz e Feijao	11
151	Semi-Arido	8
195	Amazonia Oriental	8
325	Gado de Corte	8
455	Gado de Leite	8
235	Mandioca e Fruticultura	8
624	Trigo	7
212	Hortalicas	6
542	Agropecuaria Oeste	6

L'analyse des thèmes

L'analyse des mots-clés indique, entre autres choses, les sujets préférés des divers véhicules et oriente, par ce moyen, la proposition d'articles et des suggestions de reportages par les centres de recherche. Nous avons croisé les véhicules avec les mots-clés. Notre attention est tombée sur une question polémique, les transgéniques (492) qui était le septième mot-clé le plus cité en tous les véhicules, au total, 10% du total des mots-clés. Le tableau 79 présente les détails.

Tableau 79 : Véhicules par mots-clés (transgéniques)

Total \	Véhicules	1494	1202	743	739	723	715	492
Mots-clés		producao	soja	pesquisa	projeto	tecnologia	evento	transgenicos
740	O Estado de S. Paulo	43	57	28	28	23	11	35
238	Folha de S. Paulo	16	24	10	10	14		20
505	Gazeta Mercantil	31	58	23	22	27	7	19
597	Correio do Povo	33	46	19	31	8	34	19
178	Valor Economico	10	29	13	12	8		18
128	Gazeta do Povo	9	23	7	5	2	5	18
292	Zero Hora	24	15	17	6	9	15	16
141	O Globo	5	9	10	6	8	1	16
222	O Liberal	20	8	8	18	16	8	16
165	O Povo	12	2	5	12	9	4	13
141	Jornal do Brasil	8	8	6	2	8	2	12
344	A Gazeta-MT	39	17	23	14	12	20	12
252	Correio Braziliense	9	9	16	10	13		12
555	Jornal de Brasilia	44	19	11	10	16	10	11
517	O Popular	34	37	23	11	22	9	11

Nous avons isolé le mot transgéniques (**transgenicos**) en listant les quinze véhicules qu'ont publié le plus de nouvelles caractérisées par ce mot clé. "O Estado de S. Paulo" et "Folha de S. Paulo", deux journaux nationaux se positionnent en premier. Nous avons aussi calculé le pourcentage que ces nouvelles sur les transgéniques représentaient en comparaison avec le nombre total de nouvelles publiées par le véhicule. Nous avons rencontré un grand écart, variant de 14% pour le "Gazeta do Povo" (PR) à 2% pour le "Jornal de Brasilia" (DF). Ils sont marqués en gris au tableau 80.

Tableau 80 : Véhicules par mots-clés (transgéniques) par nombre de nouvelles

Total	Mots-Clés \ Vehicules	O Estado de S. Paulo	Folha de S. Paulo	Gazeta Mercantil	Correio do Povo	Gazeta do Povo	Valor Economico	O Globo	O Liberal	Zero Hora	O Povo	Jornal do Brasil	Correio Braziliense	A Gazeta--MT	O Popular	Jornal de Brasilia
492	transgenicos	35	20	19	19	18	18	16	16	16	13	12	12	12	11	11
4895	Total	740	238	505	597	128	178	141	222	292	165	141	252	344	517	555
10	%	4,6	8,4	3,7	3,1	14	10,1	11,3	7,2	5,4	7,8	8,5	4,7	3,5	2,1	2

La prochaine étape a été d'ordonner la liste des quinze véhicules par l'importance qu'ils concèdent à ce thème, représentée par les plus grands pourcentages relatifs au nombre total de nouvelles. Là une autre surprise nous attendait. À-côté de "Gazeta do Povo" le principal journal d'un fort état agricole de la région Sud, le Parana (PR), sont apparus "O Globo", "Valor Economico", "Jornal do Brasil" et "Folha de S. Paulo", trois journaux nationaux et un spécialisé en économie. Leur intérêt par les transgéniques est deux à sept fois plus grand que celui des journaux de la queue de la liste. Ils sont marqués en gris au tableau 81.

Tableau 81 : Véhicules par mots-clés (transgéniques), importance relative des nouvelles

Total	Mots-Clés \ Vehicules	Gazeta do Povo	O Globo	Valor Economico	Jornal do Brasil	Folha de S. Paulo	O Povo	O Liberal	Zero Hora	Correio Braziliense	O Estado de S. Paulo	Gazeta Mercantil	A Gazeta--MT	Correio do Povo	O Popular	Jornal de Brasilia
États		PR	RJ	RJ	RJ	SP	CE	PA	RS	DF	SP	SP	MT	RS	GO	DF
492	transgenicos	18	16	18	12	20	13	16	16	12	35	19	12	19	11	11
4895	Total	128	141	178	141	238	165	222	292	252	740	505	344	597	517	555
10	%	14	11,3	10,1	8,5	8,4	7,8	7,2	5,4	4,7	4,6	3,7	3,5	3,1	2,1	2

Nous avons remarqué aussi que les véhicules à droite du tableau 81 étaient des états plus agricoles. La question posée était : Est-ce qu'il y a une sur valorisation du thème transgéniques par des véhicules d'un état en particulier?

Nous avons choisi de comparer la publication de dépêches des trois principaux états de la région Sud-Est, Rio de Janeiro (RJ), Sao Paulo (SP) et Minas Gerais (MG), les deux principaux états de la région Sud, Rio Grande du Sul (RS) et Parana (PR) et deux états de la région Centre-Ouest, Mato Grosso (MT) et Goiás (GO). Dans leurs territoires est produite plus de 75% de la production agricole brésilienne. Le mot-clé transgéniques est marqué en gris au tableau 82.

Tableau 82 : Véhicules par état et mots-clés

Total	Véhicules États \ Mots-clés	1	2	3	4	5	6	7	8	9	10	11	12
439	RJ	producao	transgenicos	biotecnologia	tecnologia	soja	pesquisa	projeto	clone	sementes	cultivo	gestao	biosseguranc
		35	34	22	21	19	19	17	11	11	11	9	9
5109	SP	producao	soja	tecnologia	projeto	pesquisa	leite	evento	transgenicos	agronegocio	safra	produtividade	mercado
		477	340	272	263	234	212	203	171	134	108	108	107
709	MG	producao	leite	soja	milho	cafe	transgenicos	mercado	projeto	pesquisa	sementes	frutas	tecnologia
		75	50	39	28	27	25	24	24	21	19	17	17
585	PR	soja	producao	ferrugem	pesquisa	tecnologia	transgenicos	agronegocio	safra	trigo	evento	biotecnologia	projeto
		145	44	28	28	27	25	24	22	22	20	17	16
2491	RS	producao	soja	evento	trigo	pesquisa	safra	seminario	projeto	tecnologia	mercado	transgenicos	sementes
		229	201	146	121	119	91	87	85	78	63	62	60
731	MT	producao	soja	pesquisa	evento	tecnologia	algodao	ferrugem	cultura	mercado	projeto	platio	transgenicos
		71	60	36	35	31	31	26	22	22	22	21	19
613	GO	soja	producao	pesquisa	tecnologia	algodao	leite	cerrado	cultura	frutas	contatos	safra	transgenicos
		46	42	28	26	21	20	20	19	18	17	16	16

La réponse du tableau 82 est claire. Dans la couverture de la recherche agronomique de l'Embrapa, l'état du Rio de Janeiro concède aux transgéniques un intérêt particulier, bien différente de celle d'autres états. Une observation générale sur la publication totale de nouvelles par états. Sao Paulo abrite cinq centres de recherche de l'Embrapa, Rio Grande do Sul quatre, Rio de Janeiro trois, Parana deux, Minas Gerais deux et Goias un. L'état du Mato Grosso n'en a aucun mais se poste dans une méritante troisième place quant à la quantité de nouvelles publiées sur l'Embrapa. Ceci est un indicateur d'une grande valorisation de la science, de la technologie et des innovations agricoles au Mato Grosso qui est en train de devenir d'ailleurs le principal état agricole du Brésil.

Conclusion

Les résultats démontrent avec conviction l'efficacité du système de travail de l'Embrapa pour la divulgation de ses travaux scientifiques. Ils établissent aussi des méthodologies et des outils d'analyse pour générer de l'intelligence sur les mass médias et le comportement de ses éditeurs, faire de la veille sur les médias et des audits sur le travail de communication des organisations. Ces méthodologies et outils permettent de suivre l'évolution d'un thème, d'un produit, d'une personne ou d'un secteur économique dans les médias, sur une période de temps déterminée.

La création de la base Clipping et les méthodologies et outils d'analyse ici démontrés, supposent que l'organisation recherche les informations, soit par moyen de découpages de journaux et sa digitalisation postérieure, soit par moyen de recherches sur l'Internet. Cette

opération peut être faite par des tierces parties, au dehors de l'organisation ou non, mais elle doit être coordonnée avec l'alimentation de la base de données de Clipping.

Avec des travaux systématiques d'analyse et d'audit la base Clipping peut être améliorée. Plusieurs directions se présentent. D'abord il y a un travail de re codification de variables et la réduction du nombre de véhicules suivis, surtout ces qui ont des fréquences très basses. L'analyse de la base Clipping a indiqué l'existence de 302 véhicules avec fréquence deux et un. Leur exclusion mérite d'être considéré. Après ce travail de consolidation et amélioration des variables, il y a la nécessité de veiller sur la qualité des entrées pour ne pas tomber sur la maxime d'analyser des déchets et générer des déchets. Une liste de véhicules consolidée doit être disponible. Un manuel d'indexation des données doit être fait pour guider en cas de doutes.

Les méthodologies et outils démontrés ici indiquent qu'il y a une possibilité d'automatisation en ajoutant de la richesse de détails, du travail d'analyse du Clipping, une opération qui en ce moment est achetée de tierces parties, ce qui peut représenter pour l'Embrapa aussi une importante économie de ressources financières.

L'outil VantagePoint permet d'automatiser tout le procès d'analyse par l'édition de macro commandes en langage Visual Basic, ce qui réduit le temps nécessaire pour procéder à des actualisations. Ainsi, le gérant de communication à sa disposition un portrait détaillé de la performance de son unité de travail et peut comprendre, indicateurs à l'appui, l'évolution de sa couverture dans les médias. Les bases personnalisées par centre de recherche sont aussi distribuées dans l'organisation et peuvent être vues et exploitées avec l'outil VantagePoint Reader. Ce dernier permet, avec un rapide entraînement, aux gérants et techniciens de communication manipuler les données et informations du Clipping, pour tirer leurs propres conclusions. Celle-ci est une des caractéristiques importantes des organisations intelligentes, la possibilité de disséminer le pouvoir de produire intelligence. Produire rapidement des actions plus intelligentes, contribuer pour améliorer les organisations, leurs actions et, en perfectionnant la responsabilité sociale, contribuer au développement de toute la société.

CONCLUSION GÉNÉRALE

Conclusion Générale

Les trois études de cas, du clipping électronique, des compétences de R&D et l'audit et l'analyse des bases bibliographiques de l'Embrapa sont des processus très complexes qui se composent d'un grand nombre d'étapes au sein et en dehors de l'organisation.

Il y a un point en commun dans les processus des compétences et d'audit et analyse de la production : Ce-sont des informations déjà présentes dans les organisations et qui avec des méthodologies et outils appropriés peuvent devenir "l'intelligence" organisationnelle et compétitive.

Une parcelle significative, non mesurée, des bases de données des organisations, dans le cas de l'Embrapa, pratiquement la totalité de ses bases de données, ont été créées pour recueillir et stocker des données avec la génération de rapports pré-définis pendant le processus de création et montage de la base. Ces bases ont une fonction opérationnelle importante dans les organisations. À cette fonction opérationnelle doit s'ajouter, néanmoins, une fonction d'analyse qui permet la création de systèmes d'intelligence plus performants avec l'option de réunir dans un seul environnement des ordinateurs, des données internes et externes, des méthodologies et des outils spécifiques pour cette finalité.

Audit et analyse de la production scientifique

Pour prendre en considération le fait que l'Embrapa est une organisation décentralisée, des mesures en amenant à une plus grande convergence de pratiques d'indexation concernant les documents de la Collection documentaire sont fortement recommandables. Celles-ci incluent une liste de titres des publications de la base quand un nouveau document est en train d'être indexé, une liste des journaux et revues scientifiques pour le champ "source", une liste des auteurs Embrapa pour le champ "auteur", une liste numérique des années pour le champ "date de publication" et des dictionnaires d'institutions pour un champ "affiliation". La mesure idéale pour imposer la convergence des pratiques serait l'unification des entrées dans un catalogue central en utilisant l'outil Internet et un algorithme pour la critique des nouvelles références (Hickey et Rypka, 1979) au fur et à mesure de leur catalogue.

Les résultats de l'analyse des quatre bases, Collection documentaire, Production scientifique de l'Embrapa, Pascal et WOS et leur comparaison sont aussi une indication très claire de la subjectivité des analyses de productivité de scientifiques ou d'institutions. Ces résultats ne dépendent pas seulement de la capacité des auteurs ou institutions mais surtout de la politique d'indexation des producteurs d'information. De toute manière, des indicateurs de production scientifique indiquent qui publie le plus dans la période étudiée et dans les revues scientifiques réunies dans les bases de données qui servent pour l'extraction des indicateurs.

La plainte de Velho (2001, pp.2-3) sur les indicateurs construits à partir de bases de données qui ont des vices d'origine et que pour cela ne sont pas un bon instrument de mesure pour la production scientifique d'un pays comme le Brésil nous semble justifiée, en vue de ces résultats. Ce travail pointe à des solutions pour élargir le scope des bases internationales par l'amélioration de la qualité des bases régionales et spécialisées ce qui envisagerait une insertion plus aisée de ces références par les grandes bases comme Pascal et WOS.

Les compétences de R&D

L'étude des compétences de R&D démontre avec clarté, le potentiel d'intelligence organisationnelle qui peut être généré avec des informations déjà présentes dans les organisations mais qui sont dispersées. Il démontre encore la nécessité de la création d'un environnement informatique spécifique pour l'analyse.

Pour le futur, ces informations des compétences de R&D peuvent s'intégrer à d'autres systèmes nationaux de R&D comme celui de la **Plate-forme Lattes** qui réunit des informations personnelles et professionnelles de 604.395 chercheurs, étudiants, administrateurs et d'autres personnages de la communauté scientifique Brésilienne et des pays partenaires, dont 47.973 docteurs de tous les champs de la connaissance ou le Banque de thèses de l'IBICT. Plus en avant, les méthodologies et outils ici démontrées peuvent être utilisées aussi par le CNPq et l'IBICT pour analyser leurs bases de données.

Le clipping

La fonction du clipping, d'accompagner ce qui sort dans la presse est présente dans les organisations des cinq continents. La fonction d'analyse des nouvelles est accomplie habituellement par des cabinets de conseil en médias. La création et l'alimentation de la base

Clipping et l'intégration de l'analyse au processus est le différentiel présenté par ce travail réalisé chez l'Embrapa. Avec cette base et les méthodologies et outils d'analyses ici décrites, il est possible de créer des systèmes d'intelligence et le plus important, permettre que des gérants et responsables reçoivent, manipulent, ré analysent, prennent leurs propres conclusions et surtout, agissent plus rapidement et objectivement.

La veille de l'environnement est une priorité et une nécessité des organisations pour mieux accomplir leurs missions stratégiques. L'audit du travail de communication d'une organisation soit-elle publique ou privée est un atout important pour améliorer son image et pour la construction et maintien d'un environnement favorable, pour développer une compréhension mutuelle entre l'organisation et leurs publics et contribuer ainsi à l'occurrence de débats plus éclairés sur les questions de la société.

Les mêmes méthodologies et outils peuvent aussi être appliquées dans des activités de veille des concurrents, de partenaires ou de tout un secteur. Dans l'exemple montré ici, l'objet du clipping était l'Embrapa et ses centres de recherche. Des produits et analyses similaires avec d'autres variables sont possibles avec des données du clipping d'autres organisations. La principale limitation à une application immédiate est que les bases de données de Clipping doivent être créées et maintenues. Avec des bases clippings spécifiques des applications peuvent être créées dans les champs du marketing de produit, des campagnes politiques, des audits d'image et des activités de veille de l'environnement, d'entre autres. Des études prospectives sur l'évolution historique d'un thème, d'un produit, d'un politicien ou d'une entreprise dans les médias deviennent possibles et réalisables avec une rapidité et précision qui n'est pas connue aujourd'hui dans les études médiatiques.

C'est un nouveau champ de travail dans la publicité, les relations publiques, le marketing, les campagnes politiques, qui s'ouvre pour les "bibliométristes ou infométristes", les analystes d'information versés dans les méthodologies et outils d'analyse bibliométrique et infométrique. Les débouchés sur le marché étaient limités à l'analyse de l'évolution de la science, des techniques ou technologies et des brevets.

Il est possible que le nombre total de dépêches sur les transgéniques en tous les véhicules soit beaucoup plus grand que celui enregistré par ce travail parce que cette analyse inclut seulement les dépêches de cet univers qui mentionnent l'Embrapa. Ces résultats ne

représentent pas le comportement général des véhicules. Ils reflètent le comportement des véhicules ayant référence à un thème spécifique dans la couverture de l'Embrapa.

Par utiliser une méthodologie bien fondée et robuste d'analyse bibliométrique pour des analyses des moyens de communication de masse, des bases Clipping et ces applications de médiométrie peuvent être construites et utilisées aussi dans les Écoles Supérieures ou Départements de Communication des Universités pour réaliser des analyses plus profondes sur l'action de la presse et certains aspects particuliers de leur couverture. Par exemple, la préférence des journaux de l'état du Rio de Janeiro par le thème transgéniques d'entre tous ceux de la recherche agronomique brésilienne nous fournit une pléiade de questions qui méritent d'être répondues par d'autres recherches :

D'abord, comment expliquer cette contradiction qui veut que la presse d'un état sans tradition dans l'agriculture, s'intéresse très fortement au thème transgéniques et en même temps les journaux d'états, véritablement agraires, lui donnent un traitement moins prioritaire?

Une partie de la réponse va dans le sens du côté consommateur de produits agricoles des cariocas. Autre partie de la réponse va dans le sens de la force des organisations "anti-transgéniques" au Rio de Janeiro.

Autres questions brûlantes qui méritent des études spécifiques : Est-ce que cette attitude est exclusive de la presse écrite? Est-ce que la couverture journalistique de la télévision est affectée? La rédaction de la TV Globo, le principal réseau de TV du Brésil a son siège à Rio de Janeiro. Est-ce qu'une analyse similaire faite sur les programmes journalistiques de ce réseau aura des résultats similaires?

L'étude du journalisme en sa pratique aujourd'hui dans le monde, les relations publiques et la communication d'entreprise, le marketing et aussi le segment des consultations des médias ont sur leur main des méthodologies et des outils performants qui permettent un saut de qualité et de productivité dans beaucoup d'analyses médiatiques réalisées. Des instruments pour créer des liens plus profonds et des meilleures relations entre des organisations et des sujets de toutes affiliations avec les médias et leurs éditeurs.

BIBLIOGRAPHIE

Bibliographie

ADAM, David (2002). "Citation Analysis: The Counting House". *Nature*, 415, 14 février 2002; pp.726-729.

ANGERAME, Salvador (2005). "Mesa de Negócios: Informação e Sinergia no apoio à Estratégia". In: 6ª Conferência Anual de Inteligência Competitiva. Brésil, Sao Paulo: IBC Brésil, 26-28 avril, 2005.

ANSOFF, Igor (1979). "Strategic management". États-Unis, New York, NY: John Wiley & Sons, 1979.

ARMSTRONG, Chris J. (1995). "The Eye of the Beholder". In: Basch R. (Ed.). *Electronic Information Delivery: Ensuring Quality and Value*, pp.221-244, États-Unis, Brookfield, VT: Gower Publishing Ltd., 1995. ISBN: 0-566-07567-9.

ARMSTRONG, Chris (1994). "CIQM Report on Database Quality". *Database*, décembre 1994; pp.45-48.

ASSOCIATION FRANÇAISE DE NORMALISATION-AFNOR. (1998). "XP X 50-053: Prestations de veille: Prestations de veille et prestations de mise en place d'un système de veille". Paris, avril 1998; 23p.

AYRES, F. H.; ELLIS, D.; HUGGILL, J. A. W.; YANNAKOUDAKIS, E. J. (1984). "USBC (Universal Standard Book Code): Its Use for Union File Creation; A Feasibility Study for a National Database". Angleterre, Londres: The British Library; 1984. 1 Volume (pages discontinues). ISBN: 0-7123-1020-7.

BAER, Nadine L.; BARRETT, James A.; JOHNSON, Karl E. (1995). "OPAC Database Creation Problems". *Information Technology and Libraries*, Volume 14 (3) septembre 1995; pp.179-184.

BALAJ, Sérgio T. (2005a). "Telemar Boots Sales and Satisfaction with Powerful Performance if Informatica PowerCenter". *DMReview*. septembre 2005. Disponible en <http://www.dmreview.com/article_sub.cfm?articleID=1035576>. Accès en 25/10/2005.

BALAJ, Sérgio T. (2005b). "Telemar 'Projeto Business Intelligence em DW'". In: *Inteligência Organizacional 2005 Rio de Janeiro*. Brésil, Rio de Janeiro: IDETI, 24-25 octobre 2005.

BARRÉ, Rémi (1997). "The European Perspective on S&T Indicators". *Scientometrics*, Volume 38 (1) 1997; pp.57-70.

BARROS, José R. M. de.; RIZZIERI, Juarez (2001). "Efeitos da Pesquisa Agrícola sobre o Consumidor". In: *Seminário sobre os Impactos da Mudança Tecnológica do Setor Agropecuario na Economia Brasileira*. Brésil, Brasília: Embrapa, Anales, Série Documentos n. 5, 2002.

BASCH, Reva (1995). "An Overview of Quality and Value in Information Service (Introduction)". In: Basch R. (Ed.). *Electronic Information Delivery: Ensuring Quality and Value*, pp.1-10, Brookfield, VT: Gower Publishing Ltd., 1995. ISBN: 0-566-07567-9.

BASCH, Reva (1990). "Measuring the Quality of the Data: Report on the Fourth Annual SCOUG Retreat". *Database Searcher*, Volume 6 (8) octobre 1990; pp.18-23.

BATESON, Gregory (1979). "Mind and Nature: A Necessary Unity". États-Unis, New York, NY: Bantam Books, 1979.

BAUMARD, Philippe (1991). "Stratégie et surveillance des environnements concurrentiels". France, Paris: Masson, 1991, 181p.

BELL, Daniel (1973). "The Coming Shape of Post-industrial Society, a Venture in Social Forecasting". États-Unis, New York: Basic Books, 1973, 503p.

BESSON, Bernard et POUSSIM, Jean C. (2001). "Du renseignement à l'intelligence économique". France, Paris: Dunod, 2001, 335p.

BLAKESLEE, Dorothy M.; RUMBLE JUNIOR, John (2003). "The Essentials of a Database Quality Process". *Data Science Journal*, Volume 12, février 2003; pp.35-46.

BONELLI, Regis (2002). "Impactos econômicos e sociais de longo prazo da expansão agropecuária no Brasil: revolução invisível e inclusão social". In: *Seminário sobre os Impactos da Mudança Tecnológica no Setor Agropecuario na Economia Brasileira*. Brasil, Brasília: Embrapa, *Annales, Documentos n. 5*, 2002.

BOURNE, Charles P. (1977). "Frequency and Impact of Spelling Errors in Bibliographic Data Bases". *Information Processing and Management*. Volume 13 (1) 1977; pp.1-12. ISSN: 0306-4573.

BOUTIN, Eric (2001). "A cadeia de tratamento da informação do CRRM/LePont". In: *Seminário Tecnologias para Tratamento da Informação na Embrapa*. Brasil, Brasília: Embrapa, CRRM/LePont, Cendotec, décembre 2001.

BOUTIN, Eric (1999). "Le traitement d'une information massive par l'analyse réseau : Méthode, outils et applications". Marseille: Université Aix Marseille III, Thèse (Doctorat). Laboratoire LePont, JE 196, 1999.

BRÉSIL (2002). "Decreto 4376/2002, de 13 setembro 2002, dispõe sobre a organização e o funcionamento do Sistema Brasileiro de Inteligência, instituído pela Lei nº 9.883, de 7 dezembro de 1999, e dá outras providências". Brasil, Brasília: *Diário Oficial da União*, 16 setembro 2002, rectifié le 24 septembre 2002.

BROOM, G. M.; TUCKER, K. (1989). "An essential double helix". *Public Relations Journal*, Nov. 1989, pp.39-40.

BRYANT, Patrick J.; CHU, John; HERRING, Jan P.; YOUNG, Jay (2002). "Montando uma função de inteligência competitiva: um debate em mesa-redonda". In: Prescott, J. E.; Miller, S. H. (Eds.). *Inteligência Competitiva na Prática: Estudos de casos diretamente do campo de*

batalha. Trad. Alexandre Feitosa Rosas. Brésil, Rio de Janeiro: Ed. Campus, 2002, 371p. ISBN: 85-352-0892-5.

BUENO, Wilson da C. (2002). "Medindo o retorno do trabalho de Assessoria de Imprensa". In: Duarte, J. (Org.). Assessoria de Imprensa e Relacionamento com a Mídia: Teoria e Técnica, pp.389-401. Brésil, Sao Paulo, SP: Editora Atlas, 2002, 411p.

BURTON, H. D. (1998). "Use of a Virtual Information System for Bibliometric Analysis". Information Processing and Management, Volume 24, 1998; pp.39-44. ISSN: 0306-4573.

CAHN, Pamela (1994). "Testing Database Quality". Database, Volume 17 (1) février 1994; pp.23-30.

CALLON, M.; COURTIAL, Jean P.; PENAN, H. (1993). "La Scientometrie". France, Paris: Collection Que sais-je? Presses Universitaires de France, 1993.

CANADA. FEDERAL PARTNERS IN TECHNOLOGY TRANSFER. (2001). "Links of resources for Competitive Intelligence". Disponible en < http://www.fptt-pftt.gc.ca/news/2001/compintel_e.shtml>. Accès en 28/09/2005.

CASALE, M. (1993). "CD-ROM Database Quality". Inline & CDROM Review, Volume 17 (5): octobre 1993; pp.311.

COBB, Pamela (2003). "Competitive Intelligence through Data Mining". Journal of Competitive Intelligence and Management, Volume 1 (3) Hiver 2003; pp. 80-89. Disponible en: < <http://www.scip.org/jcim.asp> >. Accès en 21/9/2005.

COELHO, Gilda M. (2005). "Qual é o Futuro da Inteligência Competitiva". In: 6ª Conferência Anual de Inteligência Competitiva. Brésil, Sao Paulo: IBC Brésil, 26-28 avril, 2005.

COELHO, Gilda M. (1999). "Sistemas de Inteligência Competitiva I". In: Curso de Especialização em Inteligência Competitiva. Brésil, Brasília: MCT/INT, CNPq/IBICT, UFRJ/ECO, 1999.

COLBY, Willian (1978). "Honorable Men: My Life in the CIA". États-Unis, New York: Simon & Schuster, 1978, 493p. ISBN: 0671228757.

COOLEY, Charles H. (1909). "The significance of communication". In: Schramm, W. & Roberts, D. (Eds.), The process and effects of mass communication, pp.643-654, 1954, États-Unis, Urbana: University of Illinois Press, 1954 Revised 1971.

CÔRTE, Adelaide R. e; ALMEIDA, Ieda M. de; ROCHA, Eulina M.; LAGO, Wilma G. do (2002). "Avaliação de softwares para bibliotecas e arquivos". Brésil, Sao Paulo: Editora Polis, 2ème édition revue et amplifiée, 2002, 221p.

COURTIAL, Jean P. (1990). "Introduction à la Scientometrie: de la bibliométrie à la veille technologique". France, Paris: Collection "Sociologies", Anthropos. 1990. ISBN: 2717819355

COYLE, Karen E.; GALLAHER-BROWN, Linda (1985). "Record Matching: An Expert Algorithm". In: Parkhurst, C. A., (Ed.). ASIS'85: Proceedings of the American Society for Information Science (ASIS) 48th Annual Meeting: Volume 22, Octobre 20-24, 1985; pp.77-80, États-Unis, Las Vegas. White Plains, NY: Knowledge Industry Publications, Inc. for ASIS; 1985. ISSN: 0044-7870; ISBN: 0-86729-176-1.

CRONIN, Blaise; DAVENPORT, Elisabeth (1993). "Social Intelligence". Annual Review of Information Science and Technology (ARIST), Volume 28 (1) 1993; pp.3-44. ISSN: 0066-4200; ISBN: 0-444-70543-0.

CUTLIP, Scott. M., CENTER, Allen. H., BROOM, Glen. M. (1985). "Effective public relations (6th ed.)". États-Unis, Englewood Cliffs, NJ: Prentice Hall, 1985, 605p.

DALESSANDRO, Glen K. (1986). "Dvorak Keyboard Implementation". In: Hurd, J. M. (Ed.) ASIS'86: Proceedings of the American Society for Information Science (ASIS) 49th Annual Meeting: Volume 23, Septembre 28-October 2, 1986; pp.59-63, États-Unis, Chicago. Medford, NJ: Learned Information Inc. for ASIS; 1986. ISSN: 0044-7870; ISBN: 0-86729-176-1.

DAMERAU, Fred J. (1964). "A Technique for Computer Detection and Correction of Spelling Errors". Communications of the ACM. Volume 5 (3) mars 1962; pp.171-176. ISSN: 0001-0782.

DEDIJER, Stevan (2003). "Development & Intelligence 2003-2053". In: Infoforum Business Intelligence Conference, Croatie, Zagreb 25-26 septembre 2003, Suède, Kund: Lund Institute of Economic Research, Working Paper Series, n. 10, 2003, 22p. ISSN: 1103-3010.

DEDIJER, Stevan (1994). "Governments, business intelligence, a pioneering report from France". Competitive Intelligence Review, Volume 5 (3) 1994; pp.45-47.

DEDIJER, Stevan (1984). "The 1984 Global System: Intelligent Systems, Development Stability and International Security". Futures, février 1984; pp.18-37.

DEDIJER, Stevan (1979a). "Breaking the monopoly on scientific know-how". Nature, Volume 280, 16 août 1979; p.529.

DEDIJER, Stevan (1979b). "Good Menus and Fine Recipes for Absent Cooks". Science, Volume 203, 1979; p.1195.

DEDIJER, Stevan (1978). "Parliaments and National Intelligence Services". Bulletin of The Atomic Scientists, Volume 34, juin 1978; pp.40-43.

DEDIJER, Stevan; JEQUIER, Nicolas (1987). "Intelligence for Economic Development: an Inquiry into the Role of Knowledge Industry". Royaume Uni, Oxford: Bergamon, 1987, 264p. ISBN: 0-85496-520-3.

DEWEY, John (1910). "How we think". États-Unis, New York, NY: Heath, 1910.

DEWEY, John (1927). "The public and its problems". États-Unis, Chicago, IL: Swallow, 1927.

- DEWEY, John (1938). "Logic: The theory of inquiry". États-Unis, New York, NY: Henry Holt, 1938.
- DEWEY, John (1939). "Theory of valuation". États-Unis, Chicago IL: University of Chicago Press, 1939.
- DOLAN, Donna R. (1992). "Quality Control at the System Level". Online, Volume 16 (2) mars 1992; pp.30-35. ISSN: 0146-5422.
- DOU, Henri (2003). "Competitive Intelligence, Trends, Methodologies and Tools". In: I Seminário Internacional Ferramentas para Inteligencia Competitiva. Brésil, Brasília: Embrapa, CRRM/LePont, 29 septembre-03 octobre, 2003.
- DOU, Henri (2001a). "Inteligência Competitiva e Gestão do Conhecimento na Embrapa, no Brasil e no mundo". In: Seminário Tecnologias para Tratamento da Informação na Embrapa. Brésil, Brasília: Embrapa, CRRM/LePont, Cendotec, décembre 2001.
- DOU, Henri (2001b). "Prefácio". In: Tarapanoff, K. (Org.) Inteligência Organizacional e Competitiva, p.27-30, Brésil, Brasília: Editora Universidade de Brasília, 2001, 343p.
- DOU, Henri (1999). "Sistemas de Inteligência Competitiva". In: Curso de Especialização em Inteligência Competitiva. Brésil, Brasília: MCT/INT, CNPq/IBICT, UFRJ/ECO, 1999.
- DOU, Henri (1995). "Veille technologique et compétitivité". France, Paris: Dunod, 1995.
- DOU, Henri (1994). "Maîtriser la prolifération de l'information grâce à la bibliométrie". Technologies Internationales, n.9, novembre 1994; pp.42-44.
- DOU, Henri (1989). "Quelques indicateurs bibliométriques en science et technique". La Tribune des Mémoires et Thèses, n.3, 1989; pp.25-28.
- DOWNHAM, Max; ALLEN, Norman J.; COSTLEY, Gary; FLYNN, Robert; GILAD, Ben; GORRIE, Thomas M.; HATHAWAY, Lorry; HERRING, Jan; MAUCH, Robert C.; PALSHO, Dorothea C.; THOMAS, David; WOOD JUNIOR, Willis (2002). "Como entender a concorrência: a perspectiva do CEO". In: Prescott, J. E.; Miller, S. H. (Eds.). Inteligência Competitiva na Prática: Estudos de casos diretamente do campo de batalha. Trad. Alexandre Feitosa Rosas. Brésil, Rio de Janeiro: Ed. Campus, 2002, 371p. ISBN: 85-352-0892-5.
- DUTHEUIL, C. (1991). "L'état de l'art de la bibliométrie en France et à l'étranger". Rapport de fin de contrat SGDN, n.24, 09 décembre 1991.
- E-CONSULTING® CORP. (2005). "Indicadores da Internet". Disponible en <http://www.econsultingcorp.com.br/insider_info/indicadores.shtml>. Accès en 17/10/2005.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA-EMBRAPA. (2006a). "Base de Dados da Pesquisa Agropecuária - BDPA". Brésil, Brasília: Disponible en <http://www21.sede.embrapa.br/bibliotecas_virtuais/index_htm>. Accès en 10/01/2006.
- EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA-EMBRAPA. Centro de Pesquisa Agropecuária dos Cerrados. (2006b). "Biblioteca Eletrônica da Embrapa Cerrados".

Brésil, Planaltina: Disponible en <http://bbeletronica.cpac.embrapa.br/versaoframe/principal_02.htm>. Accès en 10/01/2006.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA-EMBRAPA. (2005a). "Balanço Social 2004". Brésil, Brasília: Assessoria de Comunicação Social, 2005. Disponible en <<http://www22.sede.embrapa.br/publicacoes/balsoc2004/index.htm>>. Accès en 08/11/2005.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA-EMBRAPA. (2005b). "Banco de Notícias". Brésil, Brasília: Disponible en <http://www21.sede.embrapa.br/noticias/banco_de_noticias/index_htm>. Accès en 09/12/2005.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA-EMBRAPA. (2005c). "Guia de Fontes [on line]". Brésil, Brasília: Assessoria de Comunicação Social, 2005. Disponible en <http://www21.sede.embrapa.br/a_embrapa/pesquisadores/getView>. Accès en 26/07/2005.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA-EMBRAPA. (2004). "IV Plano Diretor da Embrapa 2004-2007". Brésil, Brasília: Secretaria de Gestão e Estratégica, 2004.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA-EMBRAPA. (2002). "Política de Comunicação - 2ª Edição Revista e Ampliada". Brésil, Brasília: Assessoria de Comunicação Social, 2002. Disponible en: <http://www21.sede.embrapa.br/a_embrapa/unidades_centrais/acs/publicacoes/institucional/polcomembrapa.pdf>. Accès en 23/01/2006.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA-EMBRAPA. (2001). "Manual de Informação e Documentação". Brésil, Brasília: Departamento de Tecnologia da Informação. Boletim de Comunicações Administrativas, Brasília, an 27, n. 46, 17 octobre, 2001, 96p.

EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA-EMBRAPA (1998). "III Plano Diretor da Embrapa: Realinhamento Estratégico 1999-2003". Brésil, Brasília: Secretaria de Administração Estratégica, 1998.

FARIA, Leandro I. L. (2001). "Prospecção tecnológica em materiais: aumento da eficiência do tratamento bibliométrico - uma aplicação na análise de tratamento de superfícies resistentes ao desgaste". Brésil, Sao Carlos: Universidade Federal de Sao Carlos, 2001. (Cotutelle avec l'Université de Droit, d'Economie et des Sciences d'Aix Marseille - Faculté des Sciences et Techniques de Saint Jérôme, Marseille, France) 176p, Thèse (Doctorat), 2001. Disponible en: <<http://crrm.u-3mrs.fr/res-teach/phd-defence/innocentini.html>>. Accès en 20/11/2002.

FARIA, Leandro I. L. e QUONIAM, Luc (2002). "Ferramentas para Estudos Prospectivos - Tutorial". Workshop Brasileiro de Inteligência Competitiva e Gestão do Conhecimento, 3., 2002, Brésil, Sao Paulo - SP, 16 à 18 septembre 2002, Annales.

FAUCOMPRÉ, P.; QUONIAM, Luc; DOU, Henri (1997). "The function-application relation through a link between classification and indexing". World Patent Information, v. 19, n. 3, 1997; p.167-174.

FAYARD, Pierre (2000). "O Jogo da Interação: Informação e Comunicação em Estratégia". Trad. Fachinelli, A. C., Brésil, Caxias do Sul: EDUCS. 2000, 142p.

FERREIRA, Aurélio B. de H. (1986). "Novo Dicionário da Língua Portuguesa". Brésil, Rio de Janeiro: Editora Nova Fronteira. 1ère édition. 15. impression. 1986, 1499p.

FOX, C.; LEVITIN, A.; REDMAN, T. (1994). "The Notion of Data and its Quality Dimensions". Information Processing and Management. Volume 30 (1): janvier/février 1994; pp.9-19.

FREIDSON, E. (1953). "Communication research and the concept of the mass". In: Schramm, W. & Roberts, D. (Eds.), The process and effects of mass communication, pp.197-208. États-Unis, Urbana: University of Illinois Press, 1954, Revised 1971.

FUNDAÇÃO DE AMPARO À PESQUISA DO ESTADO DE SÃO PAULO-FAPESP. (2005). "Sistema de Informações sobre Indicadores de Ciência Tecnologia e Inovação". Disponible en <<http://www.fapesp.br/indicadores/>>. Accès en 26/07/2005.

GALVIN, Robert (2002). "Inteligência competitiva na Motorola". In: Prescott, J. E.; Miller, S. H. (Eds.). Inteligência Competitiva na Prática: Estudos de casos diretamente do campo de batalha. Trad. Alexandre Feitosa Rosas. Brésil, Rio de Janeiro: Ed. Campus, 2002, 371p. ISBN: 85-352-0892-5.

GANNON, John C. (2000). "Intelligence Challenges Through 2015". Conférence au Columbus Council on World Affairs. États-Unis, 27 avril 2000. Disponible en: <http://www.cia.gov/cia/public_affairs/speeches/2000/gannon_speech>. Accès en 17/10/2005.

GASQUES, José G. (Coor.); REZENDE, Gervásio. C. de; VERDE, Carlos M. V.; CONCEIÇÃO, Junia C. P. R. da; CARVALHO, João C. de S.; SALERNO, Mario. S. (2003). "Desempenho e Crescimento do Agronegócio no Brasil". IPEA. Instituto de Pesquisa Econômica Aplicada, 2003, Brésil, Brasília - DF, octobre, 2003.

GRAHAM, Ian. S. (1997). "HTML Sourcebook". États-Unis, New York: John Wiley, 1997, 620p.

GREGOLIN, José A. R.; HOFFMANN, Wanda A. M.; FARIA, Leandro I. L.; QUONIAM, Luc; QUEYRAS, Joachim (2005). "Análise da Produção Científica a partir de Indicadores Bibliográficos". In: Landi, F. R. e Gusmão, R. (Coords). Indicadores de ciência, tecnologia e inovação em São Paulo 2004. Brésil, São Paulo: FAPESP, 2005. 2v. (992p): 02/2005. CDD 507.208161. Disponible en <<http://www.fapesp.br/indicadores>>. Accès en 27/07/2005.

GREGOLIN, José A. R.; FARIA, Leandro I. L.; HOFFMANN, Wanda A. M.; AMARAL, Roniberto M.; BARBOZA, Andréa; VALENCISE, Norberto; QUONIAM, Luc (2002). "Criação de Base de Dados para Análise Bibliométrica a partir do Conteúdo da Web of Science". 3º Workshop Brasileiro de Inteligência Competitiva e Gestão do Conhecimento, 3., 2002, Brésil, Sao Paulo - SP, 16 à 18 septembre 2002, Annales.

GRUPO DE ESTUDOS EM DOCUMENTAÇÃO-GEDOC. (SD). "Análise de Consistência de Bases de Dados". In: 1º Encontro de Aprendizagem do GEDOC. Sd. Disponible en: <<http://empresa.portoweb.com.br/ong/gedoc/estudo2.htm>>. Accès en 13/06/2005.

GRUNIG, James. E. (1992a). "Communication, public relations, and effective organizations. An overview of the book". In: Grunig, J. E. (Ed.), Excellence in public relations and communication management, pp.1-31. États-Unis, Hillsdale, NJ: Lawrence Erlbaum Associates, 1992, 666p.

GRUNIG, James. E. (1992b). "Symmetrical systems of internal communication". In: Grunig, J. E. (Ed.), Excellence in public relations and communication management, pp.531-576. États-Unis, Hillsdale, NJ: Lawrence Erlbaum Associates, 1992, 666p.

GRUNIG, James. E.; HUNT, Todd. (1984). "Managing public relations". États-Unis, Fort Worth: Holt, Rinehart and Winston, 1984, 550p. ISBN: 0030583373.

GRUNIG, James. E.; REPPER, Fred. (1992). "Strategic management, publics, and issues". In: Grunig, J. E. (Ed.), Excellence in public relations and communication management, pp.117-158. États-Unis, Hillsdale, NJ: Lawrence Erlbaum Associates, 1992, 666p.

GRUNIG, James. E.; WHITE, Jon. (1992). "The effect of worldviews on public relations theory and practice". In: Grunig, J. E. (Ed.), Excellence in public relations and communication management, pp.31-64. États-Unis, Hillsdale, NJ: Lawrence Erlbaum Associates, 1992, 666p.

GUSMÃO, Regina (2005). "Panorama Recente da CT&I em Sao Paulo: Novas Tendências, Velhos Desafios". In: Landi, F. R. e Gusmão, R. (Coords.). Indicadores de ciência, tecnologia e inovação em São Paulo 2004. Brésil, Sao Paulo: FAPESP, 2005. 2v. (992p): 02/2005. CDD 507.208161. Disponible en <<http://www.fapesp.br/indicadores>>. Accès en 27/07/2005.

HAMEL, Gary; PRAHALAD, C. K. (1994a). "Competing for the Future". États-Unis, Boston, MA: Harvard Business School Press, 1994, 327p.

HAMEL, Gary; PRAHALAD, C. K. (1994b). "Seeing the future first". Fortune, septembre 5, 1994; pp.64-70.

HERRING, Jan P. (1996). "Measuring the Effectiveness of Competitive Intelligence: Assessing & Communicating CI's Value to Organization". États-Unis, Herring & Associates and SCIP, 1996. 66p.

HICKEY, Thomas B.; RYPKA, David J. (1979). "Automatic Detection of Duplicate Monographic Records". Journal of Library Automation. Volume 12 (2) juin 1979; pp.125-142. ISSN: 0022-2240.

HIGGINS, James M.; VINCZE, Julian W. (1989). "Strategic Management: text and cases". 4^a Édition. États-Unis, New York, NY: The Dryden Press, 1989, 1180p. ISBN: 0-03-021613-3.

HOLTZ, Shel. (1999). "Public Relations on the net: winning strategies to inform and influence the media, the investment community, the government, the public, and more!" États-Unis, New York: AMA Publications, 1999.

- HOOD, William W.; WILSON, Concepcion S. (2003). "Informetric Studies using Databases: Opportunities and Challenges". *Scientometrics*. Volume 58 (3) 2003; pp.587-608.
- HUNT, C.; ZARTARIAN, V. (1990). "Le Renseignement Stratégique au Service de Votre Entreprise". France, Paris: Editions First, 1990.
- INAFUCO, Jorge K. (2005). "Convergências e Divergências: uma análise sobre a inter-relação de pesquisa e inteligência de mercado". In: 6ª Conferência Anual de Inteligência Competitiva. Brésil, São Paulo: IBC Brésil, 26-28 avril, 2005.
- INMON, William H.; HACKATHORN, Richard. D. (1994). "Using the Data Warehouse". États-Unis: Wiley - QED Publication, 1994.
- IUK RIETH GMBH. (1995). "Handbook Software Infotrans 4.0". Allemagne, Freiburg: IuK Rieth GmbH, 1995.
- JACSÓ, Peter (1997). "Content Evaluation of Databases". In: *Annual Review of Information Science and Technology (ARIST)*, Williams, M. E. (Ed.). American Society for Information Science (ASIS), Information Total Inc., États-Unis, Medford, NJ, Volume 32, 1997; pp.231-267. ISSN: 0066-4200; ISBN: 0-444-70543-0
- JACSÓ, Peter (1995). "Testing the Quality of CD-ROM Databases". In: Basch, R. (Ed.). *Electronic Information Delivery: Ensuring Quality and Value*, États-Unis, Hampshire: Grower, 1995; pp.141-168.
- JACSÓ, Peter (1993a). "Searching for Skeletons in the Database Cupboard Part I; Errors of Omission". *Database*, Volume 16(1) février 1993; pp.38-49.
- JACSÓ, Peter (1993b). "Searching for Skeletons in the Database Cupboard Part II; Errors of Comission". *Database*, Volume 16 (2) avril 1993; pp.38-49.
- JAKOBIAK, François (1991). "Pratique de la Veille Technologique". France. Paris: Les Editions d'Organisation, 1991, 232p. ISBN: 2-7081-1262-7.
- JAMBU, Michel (2000). "Introduction au Datamining: Analyse intelligente des données". France. Paris: Editions Eyrolles, 2000, 120p.
- KAPLAN, Norman (1965). "The Norms of Citation Behavior: Prolegomena to the Footnote". *American Documentation*. Volume 16 (3) juillet 1965, pp.179-184.
- KRECH, D.; CRUTCHFIELD, R. S. (1948). "Perceiving the world". In: Schramm, W. & Roberts, D. (Eds.), *The process and effects of mass communication*, pp.235-264, 1954 États-Unis, Urbana: University of Illinois Press, 1954 Revised 1971.
- LAACKMAN, Conway L.; SABAN, Kenneth; LANASA, John M. (2002). "Organização da função de inteligência competitiva: um estudo comparativo de indicadores de desempenho". In: Prescott, J. E.; Miller, S. H. (Eds.). *Inteligência Competitiva na Prática: Estudos de casos diretamente do campo de batalha*. Trad. Alexandre Feitosa Rosas. Brésil, Rio de Janeiro: Ed. Campus, 2002, 371p. ISBN: 85-352-0892-5.

- LAFOUGE, Thierry; LE COADIC, Yves F.; MICHEL, Christine (2003). "Eléments de Statistique et de Mathématique de l'Information: Infométrie, Bibliométrie, Médiométrie, Scientométrie, Muséométrie, Webométrie". France, Villeurbanne: Collection Les Cahiers de l'Enssib, Presses de l'Enssib, 2003, 319p. ISBN: 2-910227-37-5.
- LASSWELL, H. (1948). "The structure and function of communication in society". In: Schramm, W. & Roberts, D. (Eds.), *The process and effects of mass communication*, pp.84-99, 1954 État-Unis, Urbana: University of Illinois Press, 1954 Revised 1971.
- LE COADIC, Yves F. (2003). "Mathématique et Statistique en Science de l'Information: Infométrie Mathématique et Infométrie Statistique". ISDM - Information Sciences for Decision Making, France, Toulon, Volume 6 (35) 03/2003. ISSN: 1265-499X. Disponible en < http://isdsm.univ-tln.fr/articles/num_archives.htm#isdsm6>. Accès en 02/08/2005.
- LEBART, Ludovic; SALEM, André. (1994). "Statistique textuelle". France, Paris : Dunod, 1994.
- LEBART, Ludovic; SALEM, André. (1988). "Analyse statistique des données textuelles". France, Paris : Dunod, 1988.
- LEEDS, S. (2000). "Data Mining: Beware of the shaft". *Direct Marketing*. Jan. 2000. Disponible en: <<http://www.tmiassoc.com/articles/shaft.htm>>. Accès en 10/7/2002.
- LENK, Leila M.; FARIA, Ana L. D.; LIBERATTI, Ivânia A.; WALTER, Maria T. M. T. (1996). "Relatório Final: Grupo de Trabalho 'Avaliação do Aplicativo Ainfo'". Brésil, Campinas: Entreprise Brésilienne de Recherche Agronomique - Embrapa, janvier 1996. 15p.
- LENK, Leila M. (2000). "Projeto Organização, recuperação e disponibilização de informação técnico-científica em bibliotecas - Desenvolvimento do Ainfo 2000, Ainfo 2001 e Ainfo 2002". Brésil, Campinas: Entreprise Brésilienne de Recherche Agronomique - Embrapa, janvier 2000.
- LETA, Jacqueline; CRUZ, Carlos H. de B. (2003). "A produção científica Brasileira". In: Viotti, E. B. e Macedo, M. de .M. (Orgs). *Indicadores de ciência, tecnologia e inovação no Brésil*. Brésil, Campinas: Editora da Unicamp, 2003, 615p. ISBN: 85-268-0656-4.
- LEVET, Jean-Louis (2001). "L'Intelligence Economique - mode de pensée, mode d'action". France, Paris: Economica, 2001, 155p.
- LUWELL, M. (1999). "Is the Science Citation Index US-biased?" In: *Conference of the International Society for Scientometrics and Informetrics*, 7, 1999; pp.303-312. Colima, MX. *Proceedings...* États-Unis, Colima, MX: International Society for Scientometrics and Informetrics, 1999.
- LYLE, Jack; MCLEOD, Douglas (1993). "Communication media and change". États-Unis, Mountain View, CA: Mayfield, 1993.
- MACHLUP, Fritz (1983). "Semantic Quirks in Studies on Information". In Machlup, F. & Mansfield, U. (Eds.) *The Study of Information*, États-Unis, New York, NY: John Wiley, 1983; pp.641-671.

- MARCHAND, Séverine (1999). "Evaluation de la production scientifique de UCBL/INSA". Rapport de stage. Ecole Nationale Supérieure des Sciences de l'Information et des Bibliothèques - Université Claude Bernard Lyon I - France, Lyon, 1999.
- MARTRE, Henri (Coord.) (1994). "Intelligence économique et stratégie des entreprises". Commissariat Général du Plan, France, Paris: La Documentation Française, 1994, 213p.
- MASSON, Hélène (2001). "Les fondements politiques de l'intelligence économique". Thèse (Doctorat). Faculté "Jean Monnet" à Sceaux, Droit, Economie, Gestion - Université Paris Sud XI - France, Paris, 2001, 387p.
- McGONAGLE JUNIOR, J. J.; VELLA, C. M. (1996). "A new archetype for competitive intelligence". États-Unis, Westport : Quorum, 1996. 225p.
- McPHERSON, Dorothy S.; COYLE, Karen E.; MONTGOMERY, Teresa L. (1982). "Building a Merged Bibliographic Database: The University of California Experience". Information Technology and Libraries. Volume 1 (4) décembre 1982; pp.371-380. ISSN: 0730-9295.
- MEDAWAR, Katia (1995). "Database Quality: A Literature Review of the Past and a Plan for the Future". News of Computers in Libraries. Volume 29 (3) juillet 1995; pp.257-272, 1995. ISSN: 0033-0337.
- MENEZES, Diatay. B. de. (1973). "Fundamentos sociológicos da comunicação". In: Sá. A. (Comp.). Fundamentos científicos da comunicação, pp.145-205. Brésil, Petrópolis: Vozes. 1973, 499p.
- MINISTÉRIO DA AGRICULTURA, PECUÁRIA E ABASTECIMENTO-MAPA. (2005). "Balança Comercial do Agronegócio". Disponible en: <http://www.agricultura.gov.br/portal/page?_pageid=33,957256&_dad=portal&_schema=PORTAL>. Accès en 08/11/2005.
- MINISTÉRIO DA CIÊNCIA E TECNOLOGIA-MCT. (2005). "Indicadores Nacionais de Ciência & Tecnologia". Disponible en: <<http://www.mct.gov.br/estat/ascavpp/Default.htm>>. Accès en 05/8/2005.
- MINISTÉRIO DO PLANEJAMENTO, ORÇAMENTO E GESTÃO-MPO. (2000). "Curso Gestão do QPAP 2000". Brésil, Brasília: Secretaria de Gestão, 2000.
- MINTZ, Anne P. (1990). "Quality Control and the Zen of Database Production". Online, Volume 14 (6) novembre 1990; pp.14-23.
- MITCHELL, R. (1989). "Masters of innovation: How 3M keeps its new products coming". Business Week. Avril 10, 1989; pp.58-63.
- MOGEE, M. E. (1997). "Patents and technology intelligence". In: Ashton, W. B.; Klavans, R.A. (Eds.). Keeping abreast of science and technology: technical intelligence for business, États-Unis, Battelle Press, p.560, 1997.
- MORIN, Jacques (2000). "L'Excellence technologique". France : Ed. J. Picollec, 2000.

NARIN, F.; OLIVASTRO, D.; STEVENS, K. S. (1994). "Bibliometric theory, practice and problem". *Evaluation Review*, Volume 18 (1) 1994.

NOBREGA, Rosângela G. (2001). "Data Warehousing". In: Tarapanoff, K. (Org.). *Inteligência Organizacional e Competitiva*, p.285-302, Brésil, Brasília: Editora Universidade de Brasília, 2001, 343p.

NONAKA, Ikujiro; TAKEUCHI, Hirotaka (1997). "Criação de Conhecimento na Empresa: Como as Empresas Japonesas Geram a Dinâmica da Inovação". Trad. Ana Beatriz Rodrigues, Priscila Martins Celeste. Brésil, Rio de Janeiro: Campus, 1997, 358p. ISBN: 85-352-0177-7.

NORLAIN, Bernard; LA SPIERE, Loic T. (1999). "Maîtrise de l'information et prospective stratégique: l'intelligence économique au service de l'entreprise". France, Paris: Publisud, 1999, 125p.

O'NEILL, Edward T.; ROGERS, Sally A.; OSKINS, Michael W. (1993). "Characteristics of Duplicate Records in OCLC's Online Union Catalog". *Library Resources and Technical Services*. Volume 37 (1) janvier 1993; pp.59-71, 1993. ISSN: 0024-2527.

O'NEILL, Edward T.; VIZINE-GOETZ, Diane (1988). "Quality Control in Online Databases". In: Williams, M. E., (Ed.). *Annual Review of Information Science and Technology*; Volume 23. Pays Bas, Amsterdam: Elsevier Science Publishers for the American Society for Information Science; 1988. pp.125-156. ISSN: 0066-4200; ISBN: 0-444-70543-0.

OKUBO, Y. (1997). "Bibliometric indicators and analysis of research systems: methods and examples". France, Paris, OECD, 1997, 69p. (STI Working Papers, 1997/1).

OCTAVIANO, Vera L. de C. (2002). "Relatório final de auditoria realizada no Acervo Documental da Embrapa: Novembro/2001 a Fevereiro/2002". Brésil, Sao Carlos, février 2002, 79p. .

PACHECO, Carlos A. (2003). "Apresentação". In: Viotti, E. B. e Macedo, M. de .M. (Orgs). *Indicadores de ciência, tecnologia e inovação no Brasil*. Brésil, Campinas: Editora da Unicamp, 2003, 615p. ISBN: 85-268-0656-4.

PAO, Miranda L. (1989). "Importance of Quality Data for Bibliometric Research". In: Williams, M. E. (Ed.). *Proceedings of the 10th National Online Meeting*; 1989 May 9-11; États-Unis, New York, NY. Medford, NY: Learned Information Inc.; 1989. pp.321-327. ISBN: 0-938734-34-2.

PATEYRON, Emmanuel (1998). "La Veille Stratégique". France, Paris: Economica, 1998. 212p.

PENTEADO, Roberto (2005). "Análise e Mineração de textos e dados". In: Duarte, J. e Barros, A. T. de. (Orgs.). *Pesquisa em Comunicação: Métodos e Técnicas*, pp.33-48. Brésil, Sao Paulo, SP: Editora Atlas, 2005, 380p. Disponible en <http://www.editoraatlas.com.br/Atlas/uploadedFiles/1fba47c2-3637-412c-bbb8-40b76488a991Capitulos_disponiveis_site.zip>. Accès en 25/10/2005.

PENTEADO, Roberto (2002). "Assessoria de Imprensa na era digital". In: Duarte, J. (Org.). Assessoria de Imprensa e Relacionamento com a Mídia: Teoria e Técnica, pp.340-362. Brésil, Sao Paulo, SP: Editora Atlas, 2002, 411p.

PENTEADO, Roberto (2001). "Systèmes d'Intelligence Compétitive pour la Gestion Stratégique dans une Institution de Recherche et Développement". Mémoire de DEA. Laboratoire LePont, JE 196, Université de Toulon et du Var - France, Toulon, septembre 2001, 67p.

PENTEADO, Roberto (2000). "Inteligência competitiva e gestão do conhecimento na Embrapa". Projeto 14.2001.927 - SIGER - Entreprise Brésilienne de Recherche Agronomique, Unidade líder, Assessoria de Comunicação Social, Brésil, Brasília - DF, 11 de setembro de 2000.

PENTEADO, Roberto (1996). "Effects of public relations roles and models on quality committed Brazilian organizations". Mémoire Master of Arts. Department of Journalism and Communications, University of Florida - États-Unis, Gainesville, mai 1996, 161p.

PENTEADO, Roberto; QUONIAM, Luc; BOUTIN, Eric; DOU, Henri; FARIA, Leandro (2005a). "Inteligência Competitiva na Análise Estratégica das Competências: o caso Embrapa". In: 6ª Conferência Anual de Inteligência Competitiva. Brésil, Sao Paulo: IBC Brésil, 26-28 avril, 2005.

PENTEADO, Roberto; QUONIAM, Luc; BOUTIN, Eric; DOU, Henri; FARIA, Leandro (2005b). "Inteligência Competitiva na Análise Estratégica das Competências: o caso Embrapa". In: Inteligência Organizacional 2005 Brasília. Brésil, Brasília: IDETI, 15-16 juin, 2005.

PENTEADO, Roberto; DOU, Henri; BOUTIN, Eric; QUONIAM, Luc (2003a). "De la création des bases de données au développement de systèmes d'intelligence pour l'entreprise". Information Sciences for Decision Making, France, Toulon, Volume 8 (67) 05/2003; ISSN: 1265-499X, Disponible en <http://isd.m.univ-tln.fr/articles/num_archives.htm#isd.m8>. Accès en 19/09/2005.

PENTEADO, Roberto; DOU, Henri; BOUTIN, Eric; QUONIAM, Luc (2003b). "Da criação de bases de dados ao desenvolvimento de sistemas de inteligência para a organização". In: 4º Workshop Brasileiro de Inteligência Competitiva e Gestão do Conhecimento, Brésil, Salvador - BA, 20 à 22 octobre 2003.

PENTEADO, Roberto; FARIA, Leandro I. L.; VIEIRA, Job L.; KURIHARA, Maria H.; AVILA, Antonio F. D.; QUONIAM, Luc (2002). "Aplicação da bibliometria na construção de indicadores sobre a produção científica da Embrapa". In: 3º Workshop Brasileiro de Inteligência Competitiva e Gestão do Conhecimento, Brésil, São Paulo - SP, 16 à 18 septembre 2002.

PENTEADO, Roberto; QUONIAM, Luc (2001). "Aplicação da bibliometria na análise estratégica das competências da Embrapa". In: 2º Workshop Brasileiro de Inteligência Competitiva e Gestão do Conhecimento, Brésil, Florianópolis - SC, 03 à 05 octobre 2001.

PEREIRA, Josias J. de O. (2005). "Data Mining - trazendo poder ao Business Intelligence". In: Inteligência Organizacional 2005 Brasília, Brésil, Brasília - DF: IDETI, 15 à 16 juin 2005.

PEREIRA, Maria N. F.; RIBEIRO, Cláudio J. S.; TRACTENBERG, Leonel; MEDEIROS, Priscilla L. (1999). "Bases de Dados na Economia do Conhecimento: A Questão da Qualidade". *Ciência da Informação*, Volume 28 (2) mai/août 1999; pp.215-223. ISSN: 0100-1965.

PITTA, Regina A. B. (2000). "Gestion du capital humain d'une institution publique: une proposition de mapping de ressources humains". Mémoire de DEA. Faculté des Sciences et Techniques de Saint Jérôme - Université Aix Marseille III - France, Marseille, septembre 2000.

POLITY, Yves; ROSTAING, Hervé (1997). "Cartographie d'un champ de recherche à partir du corpus des thèses de doctorat soutenues pendant 20 ans: Les sciences de l'information et de la communication en France: 1974-94". In: Actes du Colloque: Les systèmes d'informations élaborées (SFBA): France, Ile Rousse, 14 a 16 juin de 1997.

PORTER, Alan L. (2003). "Text Mining for Technology Foresight". In: *Futures Research Methodology-V2.0*, The Millennium Project - American Council for the United Nations University, CÉDÉROM, 2003.

QUONIAM, Luc (2001). "Datamining, teoria e prática". In: *Seminario Tecnologias para Tratamento da Informação na Embrapa*. Brésil, Brasília: EMBRAPA, CRRM/LePont, Cendotec, décembre 2001.

QUONIAM, Luc (1999). "Datamining". In: *Curso de Especialização em Inteligência Competitiva*. Brésil, Brasília: MCT/INT, CNPq/IBICT, UFRJ/ECO, 1999.

QUONIAM, Luc (1996). "Les productions scientifiques en bibliométrie". Habilitation à diriger des recherches. Faculté des Sciences et Techniques de Saint Jérôme - Université Aix Marseille III - France, Marseille, 1996.

QUONIAM, Luc (1992). "Bibliométrie sur des références bibliographiques: méthodologie". In: Desvals, H.; Dou, H. (Org.). *La Veille Technologique*, pp.244-252. France, Paris: Dunod, 1992.

QUONIAM, Luc; HUOT, Charles; DOU, Henri; HASSANALY, Parina; BALDIT, Patrick (1993). "Bibliometric analysis of patent documents for R&D management". *Research Evaluation*, v. 3, n. 1, avril 1993; pp.13-18.

QUONIAM, Luc; BALME, François; ROSTAING, Herve; GIRAUD, Eric; DOU, Jean-Marie (1998). "Bibliometric law used for information retrieval". *Scientometrics*, Volume 41(1-2) 1998; pp.83-91.

RABAÇA, Carlos.; BARBOSA, Gustavo, (1978). "Dicionário de Comunicação com a colaboração de Muniz Sodré". Brésil, Rio de Janeiro: Editora Codecri, 1978. 512p.

REYNOLDS, George W. (1992). "Information systems for managers". États-Unis, St. Paul: West Publishing Co, 1992.

RITTBERGER, Marc; RITTBERGER, W. (1997). "Measuring Quality in the Production of Databases". *Journal of Information Science*, Volume 23 (1) 1997; pp.25-37.

ROCKART, Jean François (1979). "Chief executives define their own data needs". *Harvard Business Review*, mars, 1979; pp.81-93.

ROSENKRANS JUNIOR, Wayne A. (2002). "Passado, presente e futuro dos rumos da inteligência técnica". In: Prescott, J. E.; Miller, S. H. (Eds.). *Inteligência Competitiva na Prática: Estudos de casos diretamente do campo de batalha*. Trad. Alexandre Feitosa Rosas. Brésil, Rio de Janeiro: Ed. Campus, 2002, 371p. ISBN: 85-352-0892-5.

ROSTAING, Hervé (1993). "Veille technologique et bibliometrie: concepts, outils, applications". Thèse (Doctorat). Faculté des Sciences et Techniques de Saint Jérôme - Université Aix Marseille III - France, Marseille, 1993, 370p.

ROSTAING, Hervé (Sd.). "Guide d'utilisation de Dataview: Logiciel bibliométrique d'aide à l'élaboration d'indicateurs de tendances". Faculté des Sciences et Techniques de Saint Jérôme - Université Aix Marseille III - France, Marseille, Sd., 61p.

ROUACH, Daniel (1999). "La veille technologique et l'intelligence économique". France, Paris: Collection Que sais-je? PUF. 1999, 127p. ISBN: 2-13-047565-5.

RUSSELL, R. D. & RUSSELL, C. J. (1992). "An examination of the effects of organizational norms, organizational structure, and environmental uncertainty on entrepreneurial strategy". *Journal of Management*, Volume 18 (4) 1992; pp.639-656.

SCHRAMM, Wilbur. (1971). "How communication works". In Schramm, W. & Roberts, D. (Eds.), *The process and effects of mass communication*, pp.656-681. États-Unis, Urbana: University of Illinois Press, 1954 Revised 1971.

SEVERIN, Werner.; TANKARD JUNIOR, James. W. (1992). "Communication Theories: Origins, Methods, And Uses In The Mass Media". États-Unis, White Plains, New York: Longman Publishing Group, 1992. 364p.

SHANNON, Claude.; WEAVER, Warren. (1949). "The Mathematical Theory of Communication". États-Unis, Urbana, Illinois: University of Illinois Press, 1949.

SIEGEL, David. (1996). "Creating Killer Web Sites". États-Unis, Indiana: Hayden, 1996. 270p.

SILVA, Heloiza D. da. (2004). "Políticas de Comunicação: O Caso Embrapa". In: 1º CONVICOM - Primeiro Congresso Virtual de Comunicação Empresarial. Brésil, Sao Paulo: Contexto Comunicação e Pesquisa, juin et juillet 2004. Disponible en: "<http://www.comtexto.com.br/convicomcaseHeloizaEmbrapa.htm>". Accès en 23/01/2006.

SPINAK, Ernesto (1998). "Indicadores Cienciométricos". *Ciência da Informação*. Brésil, Brasília - DF, Volume 27 (2) 1998; pp.141-148.

STEELE, Robert D. (2005). "The Future of Intelligence: Not Federal, Not Secret, Not Expensive". In: *Speech to DHS Intelligence*, États-Unis: Washington, 10 août 2005.

Disponível em: < http://www.oss.net/extra/news/?module_instance=1&id=2633>. Acesso em 25/10/2005.

SULAIMAN, A. et SOUZA, J. M. (2001). "Data Mining Mineração de dados". In: Tarapanoff, K. (Org.) *Inteligência Organizacional e Competitiva*, p.265-278, Brasil, Brasília: Editora Universidade de Brasília, 2001, 343p.

SULE, Gisela (1986). "Problems of Duplicate Records, Standards, and Quality Control". In: Helal, A. H.; Weiss, J. W., (Eds.). *Future of Online Catalogues: Essen Symposium*; 1985 setembro 30- outubro 3; Alemanha, Essen. Essen University Library; 1986. pp.145-167. ISSN: 0721-0469; ISBN: 3-922602-09-6.

SWANSON, Don R. (2001). "ASIST Award of Merit acceptance speech: on fragmentation of knowledge, the connection explosion, and assembling other people's ideas". *Bulletin of the American Society for Information Science and Technology*, Volume 27, n. 3, 2001.

TARAPANOFF, Kira (2001). "Referencial teórico: introdução". In: Tarapanoff, K. (Org.) *Inteligência Organizacional e Competitiva*, p.13-40, Brasil, Brasília: Editora Universidade de Brasília, 2001, 343p.

TOFFLER, Alvin (1990). "Powershift. As mudanças do poder". Trad. SILVA, Luis. C. do N., Brasil, Rio de Janeiro: Editora Record, 1990. 613p.

TOFFLER, Alvin (1980). "The third wave". États-Unis, New York: Bantam Books, 1980. 539p.

TZU, Sun (1983). "A Arte da Guerra". Trad. José Sanz. 8ª edição. Brasil, Sao Paulo, SP: Record, 1983, 115p.

VAN DE VEN, A. & FERRY, D. (1980). "Measuring and assessing organizations". États-Unis, New York: Wiley, 1980.

VAN RAAN, Antony F. J. (1997). "Scientometrics: state-of-the-art". *Scientometrics*, Volume 38 (1) 1997; pp.205-218.

VELHO, Lea (2001). "Cuidado com os rankings científicos". Disponível em: <<http://www.prometeu.com.br>>. Acesso em 25/07/2005.

VILLAIN, J. (1991). "L'entreprise aux aguets". France, Paris: Editions Masson, 1991.

XIRDAL, Zéphirin (1976). "Mathématiques et sciences humaines - Union libre ou mariage forcé". *Impascience*, Volume 4 (5), Printemps 1976.

WEAVER, Warren. (1949). "Recent contributions to the mathematical theory of communication". In: Shannon, C. E. e Weaver, W. (Eds.). *The Mathematical Theory of Communication*, pp. 95-117. États-Unis, Urbana, Illinois: University of Illinois Press, 1949.

WEBSTER'S. (1994). "Webster's Encyclopedic Unabridged Dictionary of the English Language". États-Unis, Gramercy Books, 1994, 1854p.

WILKEN, E. D. (1999). "Seeking Business Intelligence in a Data Warehouse". *Graphic Arts Monthly*, Volume 71 (7) 1999; pp.71.

WILLIAMS, Martha E.; MACLAURY, Keith D. (1979). "Automatic Merging of Monographic Data Bases-Identification of Duplicate Records in Multiple Files: The IUCS Scheme". *Journal of Library Automation*. Volume 12 (2) juin 1979; pp.156-168, 1979. ISSN: 0022-2240.

WILSON, Tom D. (1988). "EQUIP: a European survey of quality criteria for the evaluation of databases". *Journal of Information Science*, Volume 24 (5) 1988; pp.345-357.

WORMELL, Irene (1998). "Informetria: explorando bases de dados como instrumentos de análise". *Ciência da Informação*, Brésil, Brasília, DF, Volume 27(2) 1998; pp.210-216.

WRIGHT, Robert (2000). "Nonzero: The Logic of Human Destiny". États-Unis, New York, NY: Pantheon Books, 2000, 448p., ISBN: 0679442529.

ANNEXES

Annexes

I. Les dix éditoriaux de l' "O Estado de S. Paulo"

Veículo	ESTADO DE S. PAULO		Data	13 Feb 2004	Quadrante
Página	A3	Fonte Citada	<input checked="" type="checkbox"/> Sem citação	<input type="checkbox"/> Dirigente	<input type="checkbox"/> Pesquisador
Composição gráfica		<input type="checkbox"/> 01 elementos gráficos	<input type="checkbox"/> 04 elementos	Presença do nome	
<input checked="" type="checkbox"/> Cabeçalho		<input type="checkbox"/> 02 elementos gráficos	<input type="checkbox"/> 05 ou mais elementos	<input type="checkbox"/> Casa	<input checked="" type="checkbox"/> Citação
Gênero	<input type="checkbox"/> Crônica	<input type="checkbox"/> Entrevista	<input type="checkbox"/> Nota Informativa	<input type="checkbox"/> Notícia	<input type="checkbox"/> Menção
<input type="checkbox"/> Artigo	<input checked="" type="checkbox"/> Escrito	<input type="checkbox"/> Carta ao leitor	<input type="checkbox"/> Nota Especial	<input type="checkbox"/> Reportagem	<input type="checkbox"/> Destaque no texto
				<input type="checkbox"/> Título	<input type="checkbox"/> Rodapé/Legenda

Emergência na Embrapa

Os sintomas são claros e alarmantes, mas ainda é possível deter o avanço do processo. Cabe ao presidente da República impedir a degeneração completa da Embrapa, uma das mais importantes instituições de pesquisa agropecuária do mundo. A atuação da empresa no jogo político partidário está minando sua organização, criando insegurança entre seus pesquisadores e pondo em risco uma orientação que foi decisiva para a modernização rural do Brasil.

Tomar a agropecuária brasileira uma das mais competitivas do mundo foi apenas uma das façanhas da Embrapa. Outro feito notável foi contribuir para o barateamento da alimentação no País, graças à elevação da produtividade agropecuária. Esse efeito, registrado nas séries de longo prazo de todas as pesquisas de custo de vida, motivou a revisão da estrutura dos vários índices de preços ao consumidor.

Essa história de sucesso, que resultou tanto em dólares para o País como em comida mais acessível aos consumidores, foi menosprezada pelo novo presidente da Embrapa, Clayton Campagnola, em sua revisão dos objetivos da empresa.

Em fevereiro do ano passado, ele definiu como "primeira vertente prioritária" as "ativida-

dades de pesquisa e desenvolvimento direcionadas aos agricultores familiares, assentados da reforma agrária e pequenos empreendedores rurais". O fortalecimento da pesquisa voltada para a expansão do agronegócio e para o aumento da competitividade virou "segunda vertente prioritária".

O presidente Campagnola negou, em entrevista publicada domingo no Estado, que a distribuição de recursos tenha mudado no primeiro ano de sua gestão ou que haja intenção de abandonar os programas em andamento. Mas o fato é que as metas prioritárias da empresa mudaram e isso deve refletir-se na sua operação nos próximos anos.

Como mostram a reportagem de Lourival Sant'Anna publicada no mesmo dia, os critérios de administração já foram alterados. Campagnola trocou 19 dos 37 chefes de centros de pesquisa e outros 9 devem ser substituídos. Alguns dos afastados tiveram o mandato interrompido. Pelo menos 10 dos 19 chefes nomeados no último ano são ligados ao PT ou ao Sindicato Nacional dos Trabalhadores de Pesquisa e Desenvolvimento Agropecuário (Sintrap).

O novo processo de seleção inclui a participação de representantes de movimentos so-

ciais nas bancas de exame. O sindicato e os movimentos influem tanto na escolha de pessoal para postos de chefia quanto na definição da política de pesquisa. Ao contrário do que disse na entrevista em resposta à reportagem de Lourival Sant'Anna, o presidente da empresa tem interferido na gestão das unidades, retirando a autonomia do chefe-geral até na escolha dos chefes-adjuntos.

O sindicato, dominado por petistas, ganhou espaço para intervir na formulação da política e para patrulhar os pesquisadores, sujeitos a execração nos comunicados da entidade quando usam, por exemplo, defensores a pesquisa de transgênicos.

A desorientação da Embrapa é visível desde o começo da gestão do atual presidente, indicado pelo ex-ministro da Segurança Alimentar José Graziano, que se notabilizou pela ineficiência na condução do Programa Fome Zero e não pelo ministro da Agricultura, como mandava o bom senso. Ao engajar-se nesse programa, a diretoria da empresa mostrou que um período de confusão e de ineficiência estava para começar.

**Política
põe em risco
uma
bem-sucedida
instituição de
pesquisa**

Pode-se medir a qualidade intelectual da nova orientação por um artigo divulgado no site da empresa no dia 18 de novembro. O texto contém preciosidades como a seguinte observação sobre as técnicas de assistência e de extensão rural: "Estas técnicas, além de dispersantes, por privilegiarem contato individual entre o técnico-

o agricultor, são alienantes por que submetem agricultor à condição de objeto do processo. No mesmo artigo, recomenda-se, como parte da nova orientação,

"incorporar o conceito de multifuncionalidade do meio rural nas ações de pesquisa e desenvolvimento, uma vez que o meio rural brasileiro não é mais exclusivamente voltado para a produção agropecuária". Entre as atividades não agrícolas lembra o autor do texto, há aquelas de turismo e lazer. Excelente lembrava, quando se trata de estabelecer uma nova orientação para a empresa de pesquisa agropecuária. É necessária uma ação de emergência para salvar a Embrapa, antes que seja tarde.

Veículo		O ESTADO DE SÃO PAULO		Data	6 E JUN 2005	Quadrante	
Página	A-3	Assunto	<input checked="" type="checkbox"/> Política <input type="checkbox"/> Agronegócio <input type="checkbox"/> Economia <input type="checkbox"/> C&T <input type="checkbox"/> Variedades	Fonte Citada	<input checked="" type="checkbox"/> Dirigente <input type="checkbox"/> Pesquisador <input type="checkbox"/> Sem citação <input type="checkbox"/> Chefe <input type="checkbox"/> Outros empregados		
Composição gráfica		<input type="checkbox"/> Texto + Foto <input type="checkbox"/> Texto + Foto + Box + Ilustração <input checked="" type="checkbox"/> Somente texto <input type="checkbox"/> Texto + Foto + Box <input type="checkbox"/> OS ou mais elementos		Presença do nome		<input type="checkbox"/> Capa <input checked="" type="checkbox"/> Citação <input type="checkbox"/> Manchete <input type="checkbox"/> Destaque no Texto <input type="checkbox"/> Título <input type="checkbox"/> Rodapé/Legenda	
Gênero		<input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input checked="" type="checkbox"/> Nota Informativa <input type="checkbox"/> Notícia <input type="checkbox"/> Artigo <input checked="" type="checkbox"/> Editorial <input type="checkbox"/> Carta ao Leitor <input type="checkbox"/> Nota Opinativa <input type="checkbox"/> Reportagem					

Má-fé e inépcia

Poucas vezes se viu, neste país, uma discrepância tão grande, entre as avaliações de um administrador público, a respeito da área de governo sob sua responsabilidade, e os fatos, eventos e problemas que então à mostra – de forma notória e escancarada, saliente-se – na realidade. Estamos falando do ministro do Desenvolvimento Agrário, Miguel Rossetto, que, indagado – em entrevista publicada em nossa edição do último domingo – se a volta do líder do MST, José Rainha Júnior, à região do Pontal do Paranapanema contribuiria para a diminuição da violência na região, respondeu que esta é sua expectativa. Ora, por tudo o que Rainha (crítico até por ex-compañheiros) já praticou, no campo da violência – ou da violência levada ao campo –, que indícios haveria de que está propenso à busca da paz na região?

A questão, no entanto, não é de erros de avaliação apenas, mas também de evidente má-fé do ministro e inépcia administrativa do governo. Alguns episódios ilustram bem a forma precipitada – para não dizer leviana – como o governo tem realizado a “sua” reforma agrária. A ministra Ellen Gracie, do Supremo Tribunal Federal (STF), concedeu liminar

que suspendeu a desapropriação, para fins de reforma agrária, da Estância do Céu, localizada no município de São Gabriel e considerada a maior já realizada no Rio Grande do Sul. Por que a suspensão do ato expropriatório assinado pelo presidente da República no mês passado? Por não ter sido respeitada a exigência legal de prévia notificação – dos proprietários ou possuidores – com indicação de dia e hora de vistoria dos imóveis a serem desapropriados. “As vistorias foram realizadas não só sem a presença do proprietário ou pessoa que pudesse prestar esclarecimentos, mas em apenas uma tarde, quando demandaria um mínimo de 15 dias” – enfatizaram os advogados de defesa dos expropriados.

E, certamente, não foi por acaso que essa “desapropriação” ocorreu no Estado natal do ministro Rossetto, onde ele, como vice-governador, apoia sistematicamente esse tipo de ação ilícita.

Por isso falamos em “má-fé” do ministro.

Outro caso: a fazendeira Maria Bethânia Vinagre acusou servidores do Instituto Nacional de Colonização e Reforma Agrária (Incra) de Belém de terem invadido a Fazenda Santa Bárbara II, em Irituia, no leste do Pará, para fazer distribui-

ção de lotes entre dezenas de famílias acampadas na área desde o fim do ano passado. E a liminar de reintegração de posse, segundo a fazendeira, até hoje não foi cumprida porque os invasores se negam a sair. Como se sabe, o ministro Rossetto distribuiu as superintendências do Incra entre ~~esses~~ aliados.

Terceiro caso: cerca de 200 pessoas, supostamente também do MST, invadiram uma área de 307 hectares, pertencente à Monsanto, em Santa Helena de Goiás (GO), destinada à pesquisa, difusão de tecnologia e beneficiamento de sementes.

Por falar em pesquisa, pior foi o caso da invasão, no dia 30 de maio, da fazenda da Embrapa (órgão do governo federal) em Ponta Grossa, Paraná, alegadamente por “suspeita” de que ali não se fizessem pesquisas e de que a propriedade tivesse sido arrendada a particulares – segundo o coordenador do MST na região, Célio Rodrigues. O diretor-executivo da Embrapa, Hebert Lima, informava, no entanto, que 2,7 mil hectares da fazenda (do total de 3,9 mil) haviam sido cedi-

dos, mas para o Instituto Agrônomo do Paraná, parceiro da Embrapa nas pesquisas.

Finalmente, temos o episódio – que o ministro considera “dramático”, mas apenas isso, como se se tratasse de fatalidade independente de quaisquer atuações governamentais de imposição de respeito à lei e à ordem – da invasão do Engenho do Prado, em Tracuateua (PE), há duas semanas, quando os sem-terra incendiaram casas (quase com moradores dentro), um galpão e várias máquinas agrícolas, além de perseguir e prender funcionários e seguranças. O ministro

afirma que escolheu, para preencher os quadros do Incra, pessoas ligadas ao MST, à Contag e à Comissão Pastoral da Terra (CPT) “porque são as pessoas mais capacitadas para desenvolver o programa (de reforma agrária)”. Nenhuma pessoa bem informada leva a sério o que diz o ministro. E o presidente da República, supostamente, é a pessoa melhor informada “deste país” sobre o que se passa no seu governo. E, felizmente, ao contrário do ministro, é um homem de boa-fé.

Ninguém bem informado pode levar a sério o que diz o ministro Rossetto

Veículo **ESTADO DE S. PAULO** Data **1 MAR 2002** 20040802

Página **3 - NOTAS** Fonte Citada Dirigente Pesquisador
 Sem citação Chefe Outros empregados

Composição gráfica 02 elementos gráficos 04 elementos
 Somente texto 03 elementos gráficos 05 ou mais elementos

Gênero Crônica Entrevista Nota Informativa Notícia
 Artigo Editorial Carta ao Leitor Nota Opínea Reportagem

Presença do nome Capa Citação
 Marchete Destaque no Texto
 Título Rodapé/Legenda

Quadrante

A	B
<input checked="" type="checkbox"/>	
C	E

Ciência pouco aplicada

O médico Paulo Marchiori Buss, presidente da Fundação Oswaldo Cruz (Fiocruz), apresentou um preocupante diagnóstico do desenvolvimento tecnológico em saúde: o conhecimento científico brasileiro progride bem, mas a aplicação prática desse conhecimento vai mal. Em entrevista publicada na edição de domingo do *Estado*, Buss lembrou que, entre 1981 e 2001, o Brasil passou de 1.889 publicações reconhecidas pelo *Institute for Scientific Information* (ISI), que é o mais relevante registro internacional de informação científica, para 10.550, um dos maiores saltos em domínio de conhecimento científico no mundo. Mas isso não se traduziu em tecnologia aplicada: entre 1990 e 1995, foram registradas no Brasil apenas 2.572 patentes – 85% delas por pesquisadores que trabalhavam fora do País. A defasagem entre teoria e prática provoca situação esdrúxula: somos, por exemplo, capazes de descobrir o princípio ativo que inibe determinada bactéria, mas não de produzir o medicamento correspondente. Buss reconheceu que só a partir de 2001 foram construídas, no instituto, as "plataformas tecnológicas capazes de finalizar o que for promissor" em pesquisas na área de vacinas e medicamentos. Como disse o presidente da Fiocruz: no Brasil, talvez apenas na agricultura a ciência não tenha ficado "só na ciência".

Esse diagnóstico – feito por quem tem autoridade e competência para fazê-lo – coloca o dedo na ferida quanto ao futuro papel da indústria brasileira. A prática da "engenharia reversa", que permitiu o desenvolvimento do Japão e do Sudoeste Asiático dos anos 60 aos 80, encerrou seu ciclo promissor. O futuro, obviamente, pertencerá a quem for capaz de juntar fronteira do conhecimento e indústria. É disto que fala o dr. Buss quando elogia a decisão do governo de ter política industrial para apenas quatro áreas, incluindo fármacos, mas reivindica que não seja esquecida, por exemplo, a biotecnologia aplicável à indústria da saúde. Ele inverte a equação tradicional e diz que "política industrial é capaz de puxar a científica".

É exatamente isso que está tentando implantar a Universidade de Campinas (Unicamp) – outro pólo nacional de desenvolvimento científico e tecnológico – quando cria uma Agência de Inovação (Inovacamp) para agilizar a ação comercial no ambiente universitário brasileiro. A função da agência é encontrar empresas dispostas a transformar as 300 patentes que são de domínio da universidade em negócio. O que a Unicamp deseja, portanto, é que a ciência que ela produz não fique "só ciência".

Esse caminho implica uma mudança de mentalidade de professores e pesquisadores, geralmente

refratários à comercialização de seus produtos. O professor João Meidanis, do Instituto de Computação da Unicamp, pioneiro do desenvolvimento da bioinformática no Brasil, revela que não foi pequena a resistência no meio acadêmico quando resolveu se transformar em um "aprendiz de empresário". Com quatro alunos, o matemático buscou financiamento e fundou a Scyla Bioinformática, única empresa brasileira na área de desenvolvimento de softwares para suporte às pesquisas nas áreas genômica, proteômica e de biologia molecular. Desde então, afirma Meidanis, em entrevista também publicada em nossa edição de domingo, "estou usando minha capacidade no lugar em que ela é mais necessária". Como ressalta o biólogo molecular Eduardo Romano, da *Embrapa*, "pesquisar e não comercializar é rasgar dinheiro público".

O futuro pertencerá a quem reunir fronteira do conhecimento e indústria

Veículo		ESTADO DE S. PAULO		Data		17 JAN 2004	
Página		Fonte Citada		<input type="checkbox"/> Dirigente <input type="checkbox"/> Pesquisador <input checked="" type="checkbox"/> Sem citação <input type="checkbox"/> Chefe <input type="checkbox"/> Outros empregados		Quadrante 	
Composição gráfica		<input type="checkbox"/> 02 elementos gráficos <input type="checkbox"/> 03 elementos gráficos <input type="checkbox"/> 04 elementos <input type="checkbox"/> 05 ou mais elementos		Presença do nome			
<input checked="" type="checkbox"/> Somente texto <input type="checkbox"/> Artigo <input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Carta ao Leitor		<input type="checkbox"/> Nota Informativa <input type="checkbox"/> Nota Colunista		<input type="checkbox"/> Capa <input type="checkbox"/> Manchete <input type="checkbox"/> Título		<input checked="" type="checkbox"/> Citação <input type="checkbox"/> Destaque no Texto <input type="checkbox"/> Rodapé/Legenda	

A reforma que Dirceu promete

O noticiário sobre o entra-e-sai de ministros deixou em segundo plano o que pode vir a ser uma reforma multissímbola mais importante do que a que se consumou na equipe do presidente e até mesmo do que a designação do titular da Casa Civil, José Dirceu, para gerente-geral da administração e vigia do desempenho do primeiro escalão do governo. Pois as novas atribuições assumidas pelo desde sempre chamado superministro de Lula, a julgar pelo teor da medida provisória que as formalizou, fazem parte do que parece constituir, ao menos no papel, um projeto de reorganização da administração federal, no mínimo tão ambicioso quanto os que o antecederam, começaram a ser postos em prática e devios estancaram, nas gestões Collor e Fernando Henrique.

O projeto concentra nas mãos de Dirceu, daquelas responsabilidades, a formidável incumbência de formular diretrizes para modernizar o Executivo, gerir os seus recursos humanos, organizar carreiras, fixar padrões de remuneração e de avaliação do trabalho dos servidores. O que isso significa efetivamente ainda é uma incógnita. Na extensa entrevista concedida ao nosso repórter Lourival Sant'Anna, publicada em nossa edição de domingo, em que falou como tendo já incorporado as funções de um ministro da Administração,

Dirceu disse que a meta é estender para todo o governo o padrão de excelência de organismos como "o Banco Central, a Receita, o Itamaraty, as Forças Armadas, a estrutura do Planalto" e de ministérios como "o da Saúde e Educação, que têm corpo burocrático, políticas públicas, eficiência".

"Uma das principais tarefas que tenho agora", informou, em dado momento, "é fazer mudanças no Inbra, no Ibama, na Embrapa, na Funai...". É estranho e preocupante a inclusão da Embrapa nessa heterogênea coleção. Uma das ilhas de excelência no arquipélago das estatais brasileiras, a única mudança a que a empresa aspira é um orçamento que lhe permita expandir as pesquisas em agropecuária pelas quais se tornou conhecida no mundo e manter um corpo de cientistas à altura dos seus objetivos. É de esperar que a referência de Dirceu à Embrapa seja antes um lapso do que um indicio de que a pretendida reorganização da administração direta e indireta federal desembocará na proverbial emenda pior do que o soneto — ou, de novo, na reinvensão da roda.

Por isso, o governo deve deixar claro o quanto antes, para além das generalidades sobre a

Veículo	ESTADO DE S. PAULO 2		Data	27 JAN 2004	Quadrante <table border="1"> <tr> <td>A</td> <td>B</td> </tr> <tr> <td>C</td> <td>D</td> </tr> <tr> <td></td> <td>E</td> </tr> </table>	A	B	C	D		E
A	B										
C	D										
	E										
Página	3-A		Fonte Citada	<input type="checkbox"/> Dirigente <input type="checkbox"/> Pesquisador <input type="checkbox"/> Sem citação <input type="checkbox"/> Chefe <input type="checkbox"/> Outros empregados							
Composição gráfica	<input type="checkbox"/> 02 elementos gráficos <input type="checkbox"/> Somente texto	<input type="checkbox"/> 03 elementos gráficos <input type="checkbox"/> 04 elementos <input type="checkbox"/> 05 ou mais elementos	Presença do nome								
Gênero	<input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Carta ao Leitor	<input type="checkbox"/> Nota Informativa <input type="checkbox"/> Nota Operativa	<input type="checkbox"/> Notícia <input type="checkbox"/> Reportagem	<input type="checkbox"/> Capa <input type="checkbox"/> Manchete <input type="checkbox"/> Título	<input type="checkbox"/> Citação <input type="checkbox"/> Destaque no Texto <input type="checkbox"/> Rodapé/Legenda						

busca da eficiência na estrutura federal e do aperfeiçoamento dos mecanismos anticorrupção no seu interior, se a reforma fará o mais necessário - dotar o Estado de uma burocracia profissionalizada. Um quadro permanente de servidores de elite é essencial para melhorar a qualidade da administração, assegurar a sua continuidade e pôr a salvo do toma-lá, dá-cá a gestão da coisa pública - que

exige técnicos competentes e não apadrinhados de políticos influentes. A regra geral é singela: à parte quaisquer outros fatores, o setor público será tão mais eficiente quanto menos afetado pelo vaivém dos governantes e dos partidos no poder.

Na França, para citar um exemplo clássico, a Quarta República (de 1945 a 1958) foi um sem-fim de governos cuja sobrevida se media em meses. Mas esse paroxismo de instabilidade política em nada afetou a estabilidade administrativa - fosse quem fosse o primeiro-ministro de turno, o aparelho estatal seguia funcionando normalmente. A propósito, a palavra *establishment*, usada hoje como sinônimo do conjunto das elites de um país, foi cunhada pelo autor inglês Anthony Sampson, em 1962, pa-

ra definir a burocracia britânica, que permaneceu estável apesar do até então inédito revezamento de governos conservadores e trabalhistas, à época ferozmente antagônicos, a partir do segundo pós-guerra.

Nesses países, bem como nos Estados Unidos, cada novo chefe de governo pode nomear umas poucas centenas, se tanto, de autoridades e burocratas - em vez dos muitos milhares

O País precisa de uma burocracia profissional qualificada e permanente

de "cargos de confiança" à disposição dos presidentes brasileiros para o escambo político. Ainda agora, o governo abriu 2.797 novas vagas no Executivo

- das quais nada menos de 1.332 são daquele tipo. O que também leva a perguntar se o que se quer é modernizar a máquina ou "aparelhá-la", para mantê-la sob controle do PT mesmo depois que Lula deixar o Planalto. E, se a resposta for mesmo modernização, o que só mereceria aplauso, o ministro Dirceu deverá demonstrar, na frente administrativa, a mesma competência que lhe é reconhecida na esfera política - porque enorme é o seu desafio e enormes são os poderes que passou a acumular.

Veículo O ESTADO DE S.PAULO		Data 13 ABR 2003	Quadrante						
Página Q8.f	Assunto <input type="checkbox"/> Política <input type="checkbox"/> Agronegócios <input type="checkbox"/> Economia <input type="checkbox"/> C&T <input checked="" type="checkbox"/> Variedades	Fonte Citada <input type="checkbox"/> Dirigente <input type="checkbox"/> Pesquisador <input type="checkbox"/> Sem citação <input type="checkbox"/> Chefe <input type="checkbox"/> Outros empregados	<table border="1"> <tr><td>A</td><td>B</td></tr> <tr><td>C</td><td>D</td></tr> <tr><td>E</td><td></td></tr> </table>	A	B	C	D	E	
A	B								
C	D								
E									
Composição gráfica <input checked="" type="checkbox"/> Somente texto <input type="checkbox"/> Texto + Foto <input type="checkbox"/> Texto + Foto + Box <input type="checkbox"/> Texto + Foto + Box + Ilustração <input type="checkbox"/> Os ou mais elementos		Presença do nome <input type="checkbox"/> Cabe <input type="checkbox"/> Citação <input type="checkbox"/> Manchete <input type="checkbox"/> Destaque no Texto <input checked="" type="checkbox"/> Título <input type="checkbox"/> Rodapé/Legenda							
Gênero <input type="checkbox"/> Artigo <input checked="" type="checkbox"/> Editorial <input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Carta ao Leitor <input type="checkbox"/> Nota Informativa <input type="checkbox"/> Nota Opinativa <input type="checkbox"/> Notícia <input type="checkbox"/> Reportagem									

É preciso preservar a Embrapa

Houve uma revolução silenciosa na economia brasileira, nos anos 80 e em boa parte dos 90. Essa revolução tornou mais fácil e menos dolorosa a política de estabilização iniciada com o Plano Real. A grande mudança foi a redução do preço real dos alimentos, que permitiu a milhões de brasileiros, na última década, melhores padrões de nutrição. Entre 1975 e o ano 2000, o preço de uma cesta composta de leite, carne bovina, frango, arroz, feijão, ovos, café e outros 11 produtos caiu em média 5,25% ao ano, segundo cálculos de pesquisadores da USP. Essa mudança resultou de enormes ganhos de produtividade permitidos pela mudança tecnológica. Por trás dessa história está uma das instituições mais sérias e mais eficientes do País, a Empresa Brasileira de Pesquisa Agropecuária (Embrapa). A crise financeira que essa organização vem enfrentando – pela segunda vez desde o começo dos anos 90 – é, portanto, um grave motivo de preocupação para quem se interessa pela modernização do Brasil e pelo bem-estar da maior parte das famílias brasileiras.

Neste ano, a Embrapa recebeu pouco mais de 3% das verbas previstas para investimento e custeio de suas pesquisas, embora já se tenha completado um quarto do exercício fiscal. Algumas de suas unidades não têm dinheiro para pagar as contas de luz e telefone. As linhas telefônicas de pelo menos uma delas – a de Rondônia – já não funcionam, segundo informou o Estado na edição de quinta-feira. A cúpula da empresa procura ser moderada ao descrever a situação, lembrando que há um novo governo e que toda a liberação de verbas do Tesouro tem sido retardada. Não nega, no entanto, que as verbas venham encolhendo e que sejam insuficientes para as atividades da instituição.

O corte de verbas e o atraso na liberação de recursos afetam de forma diferente cada área do setor público. Os danos podem ser especialmente grandes, e até irreversíveis, numa instituição como a Embrapa. Alguns projetos de pesquisa agropecuária consomem muito tempo – anos, em muitos casos – e não se pode simplesmente interrompê-los. A interrupção, em alguns casos, pode resultar na perda de todo o trabalho realizado até uma data. O prejuízo é múltiplo. Perdem-se o dinheiro investido no estudo, o tempo dedicado ao trabalho e, naturalmente, os benefícios econômicos e sociais que surgiram com a aplicação

dos conhecimentos adquiridos.

No caso da pesquisa agropecuária, esses benefícios são muito grandes e seu alcance nem sempre é percebido. Quando a alimentação se torna mais barata, como ocorreu nas duas últimas décadas, uma parte do orçamento familiar é liberada para poupança ou para outras despesas. As indústrias de bens duráveis e semiduráveis também são beneficiadas, portanto, pelo aumento da produtividade agrícola. Os benefícios da pesquisa vão muito além do suprimento de comida para a mesa das famílias brasileiras.

As pesquisas têm permitido, também, uma enorme ampliação de oportunidades econômicas. Neste ano, o Brasil deve faturar mais que os Estados Unidos com a exportação de soja e derivados. Há 30 anos, a cultura da soja, uma planta originária de climas temperados, ainda estava começando na Região Sul do Brasil. Hoje, essa cultura cobre grandes áreas também nas zonas tropicais, graças ao trabalho de adaptação realizado pela Embrapa. Da mesma forma, o avanço tecnológico tornou o frango uma fonte barata e abundante de proteína para os brasileiros. Esse mesmo frango, além disso, é um importante componente da pauta de exportações.

É difícil, em tempos de aperto fiscal, evitar que atividades importantes sejam prejudicadas pelo corte de verbas ou pelo atraso na liberação de recursos. É um dos preços do ajuste das contas públicas. Mesmo assim, é sempre possível – e necessário – levar em conta as peculiaridades de cada setor da administração, na hora de programar os cortes. Há excelentes motivos para que a pesquisa agropecuária seja tratada, nessas ocasiões, com muito cuidado.

"Nenhum governo pode se dar ao luxo de chegar ao fim do ano e não ter gasto o que está determinado no orçamento das instituições." Essa declaração, publicada na mesma edição do Estado em que saiu a reportagem sobre a crise da Embrapa, foi feita pelo presidente Lula, criticando o governo Fernando Henrique por ter feito isso com o orçamento para saneamento e infra-estrutura.

A pesquisa agropecuária não ajuda só o campo, mas toda a economia

Veículo **A ESTÂNCIA DE A. PAULO** Data **5 0 ABH** Quadrante

Página 23	Assunto <input type="checkbox"/> Economia <input type="checkbox"/> Política <input type="checkbox"/> Agronegócio <input type="checkbox"/> C&T <input type="checkbox"/> Variações	Fonte Citada <input type="checkbox"/> Dirigente <input type="checkbox"/> Pesquisador <input type="checkbox"/> Sem citação <input type="checkbox"/> Chefe <input type="checkbox"/> Outros empregados	<table border="1"> <tr> <td>A</td> <td>B</td> </tr> <tr> <td>C</td> <td>D</td> </tr> <tr> <td>E</td> <td></td> </tr> </table>	A	B	C	D	E	
A	B								
C	D								
E									
Composição gráfica <input checked="" type="checkbox"/> Somente texto <input type="checkbox"/> Texto + Foto <input type="checkbox"/> Texto + Foto + Box <input type="checkbox"/> Texto + Foto + Box + Nustração <input type="checkbox"/> OS ou mais elementos	Presença do nome <input type="checkbox"/> Capa <input type="checkbox"/> Manchete <input type="checkbox"/> Título <input type="checkbox"/> Criação <input type="checkbox"/> Destaque no Texto <input type="checkbox"/> Rodapé/Legenda								
Gênero <input type="checkbox"/> Artigo <input checked="" type="checkbox"/> Editorial <input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Carta ao Leitor <input type="checkbox"/> Nota Informativa <input type="checkbox"/> Nota Opinativa <input type="checkbox"/> Notícia <input type="checkbox"/> Reportagem									

Um pedaço brasileiro de Primeiro Mundo

Até domingo Ribeirão Preto será a capital latino-americana da tecnologia agrícola. Quem visitar o Agrishow, a maior feira do agronegócio ao sul do Rio Grande, conhecerá um pedaço de Primeiro Mundo em território brasileiro: um setor que se preparou para enfrentar a competição internacional e que garante mais de metade da receita nacional de exportações. Algumas indústrias brasileiras progrediram muito desde os anos 80. Mas em nenhum setor de negócios a modernização e os ganhos de eficiência foram tão amplos e tão consistentes quanto na agropecuária. A feira de Ribeirão Preto, que se realiza anualmente desde 1993, ajuda a entender como se deu essa metamorfose.

Nos últimos 13 anos, a área cultivada com lavouras de grãos e oleaginosas cresceu apenas 12%. No mesmo período, a produção dessas culturas aumentou 93%, como lembrou o ministro da Agricultura, Roberto Rodrigues, na inauguração da feira deste ano. Essa evolução foi o efeito de vários fatores combinados - pesquisa tecnológica, difusão de novas práticas, muito investimento em meios de produção e uma pilada de políticas sensatas. O

financiamento de máquinas e implementos pelo Moderfrota, um programa do Banco Nacional do Desenvolvimento Econômico e Social (BNDES), lançado há mais de dois anos, foi uma dessas políticas.

Os ganhos de produtividade, no entanto, começaram muito antes, em geral sob a liderança de alguns empresários inovadores, que mostraram aos demais as vantagens da mudança. As transformações foram realizadas, até há pouco tempo, em condições adversas. Desde o começo dos anos 90, o financiamento oficial à agricultura escasseou. Subsídios foram cortados, o crédito encolheu em termos reais e os programas de comercialização foram em grande parte abandonados, porque o Tesouro se tornou incapaz de manter as políticas tradicionais.

Novas formas de financiamento foram buscadas, parte nos mecanismos oficiais, parte no sistema privado. A Embrapa e outras instituições de pesquisa continuaram, apesar das dificuldades, a contribuir para a mudança tecnológica. Esse apoio foi precioso não só para o aumento da produção por hectare, mas também - mais dramaticamente - para a recuperação de uma

cultura, a do algodão, devastada pela concorrência do produto importado.

A transformação da agropecuária foi essencial, na última década, para a melhora do abastecimento interno e para a geração de dólares por meio do comércio exterior. Ao reduzir o custo dos alimentos, esse ganho de eficiência facilitou, internamente, mais que o controle da inflação. Permitiu que as famílias desviassem uma parcela maior de seu orçamento à compra de bens duráveis e semiduráveis, elevando seu padrão de vida e ampliando o mercado da indústria. É preciso levar em conta esses dados, nem sempre percebidos, para bem avaliar o peso econômico do agronegócio.

O ministro da Agricultura, Roberto Rodrigues, anunciou em Ribeirão Preto a criação de um programa de renda mínima para o produtor rural. O assunto ainda está em estudo, em Brasília, mas é evidente que os beneficiários serão principalmente os pequenos produtores.

Mas a expressão "pequeno produtor" designa agriculto-

res que operam em condições muito diferenciadas em todo o Brasil. Alguns deles são tecnicamente bem preparados e já participam com eficiência da atividade comercial. Muitos têm contratos de fornecimento para indústrias.

Há uma enorme distância - econômica, tecnológica e cultural - entre esses e outros pequenos produtores, muitos dos quais só conseguem sobreviver em condições primitivas. Qualquer programa de apoio à agricultura familiar, incluída a reforma agrária, só será eficaz se essas diferenças forem

consideradas com realismo.

Os padrões de referência para a produtividade física e a eficiência econômica são aqueles encontrados na feira de Ribeirão Preto. São as condições necessárias à sobrevivência em mercados competitivos. Qualquer política orientada por padrões mais baixos tenderá a converter-se em assistencialismo, provavelmente custoso e ineficiente a longo prazo. Esse é um bom tema de meditação que uma visita ao Agrishow pode proporcionar.

O Agrishow mostra um Brasil que dá certo e ensina o caminho da modernização

Veículo		O ESTADO DE S.PAULO		Data		11 MAI 2002		Quadrante	
Página		Assunto		Fonte Citada		Dirigente		Pesquisador	
3 - MATAS		<input checked="" type="checkbox"/> Política <input type="checkbox"/> Agronegócio <input type="checkbox"/> Economia <input type="checkbox"/> C&T <input type="checkbox"/> Variedades		<input checked="" type="checkbox"/> Sem citação <input type="checkbox"/> Chefe		<input type="checkbox"/> Outros empregados			
Composição gráfica		<input type="checkbox"/> Texto + Foto <input type="checkbox"/> Texto + Foto + Box + Ilustração <input checked="" type="checkbox"/> Somente texto <input type="checkbox"/> Texto + Foto + Box <input type="checkbox"/> OS ou mais elementos		Presença do nome		<input type="checkbox"/> Capa <input type="checkbox"/> Citação <input type="checkbox"/> Manchete <input checked="" type="checkbox"/> Destaque no Texto <input type="checkbox"/> Título <input type="checkbox"/> Rodapé/Legenda			
Gênero		<input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Nota Informativa <input type="checkbox"/> Notícia <input type="checkbox"/> Artigo <input checked="" type="checkbox"/> Editorial <input type="checkbox"/> Carta ao Leitor <input type="checkbox"/> Nota Opinitiva <input type="checkbox"/> Reportagem							

Boa safra e boas exportações

Agropecuária brasileira continua a mostrar excelente desempenho tanto na produção quando no comércio exterior. Recordes de exportação e de saldo comercial foram anunciados na quarta-feira. Na quinta, o ministro da Agricultura, Roberto Rodrigues, divulgou a nova estimativa da safra de grãos e oleaginosas, que poderá chegar a 115,12 milhões de toneladas. Se esse total for confirmado, a produção da temporada 2002/2003 será 19,1% maior que a anterior. O ministro já se arisca a prever um aumento de 5% para a próxima colheita, que poderá totalizar 120 milhões de toneladas.

O desempenho comercial continua a melhorar de ano para ano. O agronegócio brasileiro – que inclui matérias-primas e alguns produtos elaborados e semi-elaborados – faturou US\$ 8,19 bilhões com a exportação, de janeiro a abril deste ano, e conseguiu um superávit de US\$ 6,63 bilhões. O valor exportado foi 31,1% maior que o de um ano antes. O superávit ficou 40,7% acima do contabilizado nos primeiros quatro meses de 2002.

As vendas ao exterior cresceram em volume e foram favorecidas pela recuperação de alguns preços, como o da soja. A receita de soja e

derivados foi 87,9% maior que a de janeiro-abril do ano passado. A de papel e celulose, 43,7%. A de carnes, 23%. Os exportadores elevaram as vendas tanto para mercados tradicionais, como a União Européia (26,9%), quanto para novos parceiros, como a China (quase 400%).

A eficiência dos produtores brasileiros cresceu acentuadamente nos últimos dez anos. Isso explica boa parte do aumento das vendas ao exterior. Desde a mudança cambial de janeiro de 1999, agricultores e pecuaristas puderam aproveitar mais amplamente, no mercado internacional, a produtividade conquistada nos anos anteriores. Pelos cálculos do ministério, a produção de grãos e oleaginosas cresceu 99% em 12 anos, ou 5,9% ao ano. A área plantada aumentou apenas 12,7% nesse período, cerca de 1% ao ano.

Isso se explica pelos avanços tecnológicos, como sementes melhoradas, fertilizantes, defensivos e melhor manejo de solo. Na produção animal, as grandes mudanças começaram antes. Graças a isso aumentaram simultanea-

mente o consumo interno de carnes, ovos e laticínios e a exportação de produtos de origem animal. De janeiro a abril deste ano, por exemplo, o País faturou US\$ 1,02 bilhão com a exportação de carnes, 23% mais que em igual período de 2002.

Os ganhos de produtividade continuam presentes na safra da temporada 2002/2003. Os bons resultados obtidos nos últimos anos estimularam aumento de

10,8% na área plantada com soja. A produção, no entanto, deve ter crescido 20,1%, segundo o último levantamento da Companhia Nacional de Abastecimento (Conab). A safra deve chegar a 50,3 milhões de toneladas. Os produtores colheram, até agora, cerca de 90% desse volume. A produção de milho deve ter alcançado 33,7 milhões de toneladas na safra de verão, 15,8% maior que a de um ano antes. É o resultado da normalização do tempo, já que em 2002 a safra desse produto foi prejudicada pela estiagem.

Com a segunda safra, ou safrinha, a produção total de milho poderá atingir o recorde de 42,8 milhões de toneladas. Esse é um insumo fundamental para a produção de aves e suínos, destinados tanto ao mercado interno quanto à exportação.

O agronegócio continua, portanto, a cumprir com êxito um duplo papel: garantir o abastecimento interno – em condições cada vez melhores – e proporcionar ao Brasil um robusto faturamento em dólares.

O setor continuará a melhorar seu desempe-

Agropecuária mostra como crescer e competir sem recorrer ao protecionismo

nho, de enorme importância para a estabilidade econômica do País, se algumas condições simples forem observadas pelo governo. Os produtores precisam de financiamento a custo razoável e na hora certa, de acesso a tecnologia atualizada – em boa parte proporcionada pela Embrapa – e de tranquilidade para trabalhar. Além disso, o governo ajudará o agronegócio se for capaz, na área internacional, de conseguir condições equitativas de competição, com menos barreiras comerciais e menos distorções dos mercados.

Veículo		ESTADO DE S.PAULO		Data		250 98 03		Quadrante							
Página		Assunto		Fonte Citada		Presença do nome		<table border="1"> <tr> <td>A</td> <td>B</td> </tr> <tr> <td>C</td> <td>D</td> </tr> <tr> <td>E</td> <td></td> </tr> </table>		A	B	C	D	E	
A	B														
C	D														
E															
<input checked="" type="checkbox"/> 3 A		<input type="checkbox"/> Política <input type="checkbox"/> Agropecuária <input type="checkbox"/> Economia <input type="checkbox"/> C&T <input checked="" type="checkbox"/> Variedades		<input type="checkbox"/> Sem citação <input type="checkbox"/> Dirigente <input type="checkbox"/> Chefe <input checked="" type="checkbox"/> Pesquisador <input type="checkbox"/> Outros empregados		<input type="checkbox"/> Capa <input type="checkbox"/> Citação <input type="checkbox"/> Manchete <input checked="" type="checkbox"/> Destaque no Texto <input type="checkbox"/> Título <input type="checkbox"/> Rodapé/Legenda									
Composição gráfica		<input type="checkbox"/> Texto + Foto <input type="checkbox"/> Texto + Foto + Box + Ilustração <input type="checkbox"/> OS ou mais elementos													
<input checked="" type="checkbox"/> Somente texto <input type="checkbox"/> Texto + Foto + Box															
Gênero		<input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Nota Informativa <input type="checkbox"/> Nota <input type="checkbox"/> Artigo <input checked="" type="checkbox"/> Editorial <input type="checkbox"/> Carta ao Leitor <input type="checkbox"/> Nota Opinativa <input type="checkbox"/> Reportagem													

Em defesa dos transgênicos

Se o Planalto cumprir o que prometeu em meados de maio, a fim de assegurar a aprovação da medida provisória que liberou a soja transgênica plantada ilegalmente no ano passado, deverá chegar ao Congresso até a próxima semana o projeto de lei, a tramitar em regime de urgência, que definirá a política da administração Lula em relação ao plantio de espécies que contenham organismos geneticamente modificados (OGM).

Apesar das pressões encarniçadas de setores que desejam impedir que os agricultores brasileiros tenham acesso aos recursos da biotecnologia – sem oferecer um único argumento objetivo, capaz de respaldar do ponto de vista científico e econômico essa obstinada recusa –, seria um verdadeiro choque se as normas que se presume estejam sendo elaboradas pelo governo equivalassem, afinal, a uma “proibição branca” da agricultura transgênica.

Para tanto, bastaria sujeitar os pedidos de cultivo de OGM a um emaranhado de exigências burocráticas praticamente impossíveis de serem cumpridas e conceder a última palavra nessas solicitações exatamente àqueles setores que, deturpando o princípio da precaução pelo qual o poder público deve se orientar nessa matéria, gostariam de impor aos transgênicos uma moratória não declarada por tempo indeterminado.

Esse veto que não ousa dizer o nome já vem travando as pesquisas em biotecnologia, com claros prejuízos para a capacitação do Brasil numa área de inovação científica de ponta, destinada a um desenvolvimento notável neste século. A extravagante equiparação a “agrotóxicos e afins” de plantas alteradas para resistir a doenças e insetos é um exemplo dessa esperteza.

Os pesquisadores interessados em trabalhar com tais organismos precisam do mesmo Registro Especial Temporário (RET) para venenos químicos fornecido pelo Ibama. Além da papelada necessária, o Ibama frequentemente não se manifesta no prazo legal de 60 dias. Só na **Embrapa** há projetos parados há três anos. “Ou o pesquisador trabalha fora da lei ou tem de abandonar os projetos”, resume um estudioso.

“Trabalhar fora da lei” é uma alternativa admitida abertamente por não raros sojicultores do Rio Grande do Sul, na implausível hipótese de o governo deixar de legalizar esse cultivo, tamanhas as suas vantagens econômicas. “O custo é menor e o retorno, maior”, explica um produtor gaúcho, citado pela *Gazeta Mercantil*, referindo-se à economia no preço do

2600653

Veículo		O ESTADO DE S.PAULO		2	Data	10 JUN 2003	Quadrante
Página	Assunto	<input type="checkbox"/> Política	<input type="checkbox"/> Agronegócio	Fonte Citada	<input type="checkbox"/> Dirigente	<input type="checkbox"/> Pesquisador	
C3 A	<input type="checkbox"/> Economia	<input type="checkbox"/> C&T	<input checked="" type="checkbox"/> Variedades	<input type="checkbox"/> Sem citação	<input type="checkbox"/> Chefe	<input type="checkbox"/> Outras empregados	
Composição gráfica		<input type="checkbox"/> Texto + Foto		<input type="checkbox"/> Texto + Foto + Box + Ilustração		Presença do nome	
<input checked="" type="checkbox"/> Somente texto		<input type="checkbox"/> Texto + Foto + Box		<input type="checkbox"/> OS ou mais elementos		<input type="checkbox"/> Capa	
Gênero		<input type="checkbox"/> Crônica		<input type="checkbox"/> Entrevista		<input type="checkbox"/> Nota Informativa	
<input type="checkbox"/> Artigo		<input type="checkbox"/> Editorial		<input type="checkbox"/> Carta ao Leitor		<input type="checkbox"/> Nota Opinativa	
		<input type="checkbox"/> Notícia		<input type="checkbox"/> Reportagem		<input type="checkbox"/> Manchete	
		<input type="checkbox"/> Destaque no Texto		<input type="checkbox"/> Rodapé/Legenda		<input type="checkbox"/> Título	

herbicida utilizado e aos ganhos de produtividade da lavoura.

Outro produtor observa que "a maioria dos plantadores multiplica sementes dentro da própria fazenda". Haveria cerca de 30 variedades de sementes OGM sendo vendidas ilegalmente do Sul ao Nordeste, passando pelo Sudeste e o Centro-Oeste. A própria **Embrapa Soja** tem perto de 20 sementes distintas, à espera da libera-

ção legal, para uso "nos quatro cantos do País", informa o jornal.

Diante disso, e incapazes de provar que os transgênicos são mais perigosos para a saúde e o ambiente do que os produtos convencionais - na realidade, tendem a ser menos perigosos, pela redução no volume de inseticidas e herbicidas utilizados nas plantações -, os ecofundamentalistas apelam para a economia: sendo o único grande produtor mundial de soja não transgênica, o Brasil terá mercados cativos nos países refratários aos OGM, sobretudo na Europa.

É falso. Embora a agricultura transgênica esteja suspensa na União Européia, o consumo desses alimentos é livre, e as exportações de soja dos Estados Unidos e da Argentina

(no caso, toda ela GM) para a região aumentaram 50% no último ano. A UE importa mais de 41 milhões de toneladas anuais de soja, em grão ou como farelo. De qualquer forma, o prêmio pago pelo produto tradicional é menor do que o lucro proporcionado ao plantador brasileiro pela soja GM.

Os europeus devem mudar a sua política para os transgênicos, acredita o cientista Inge

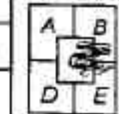
Russell, da Universidade de Edimburgo, ouvido pelo jornal *Valor*. Em consequência, raciocina, mesmo que o Brasil perca algum mercado no curto prazo

ao liberar os transgênicos, "correrá o risco de perder um mercado ainda maior se não estiver no jogo". E quando estiver - as opiniões são unânimes - simplesmente não haverá soja convencional para atender à demanda da UE. A força das coisas levará os europeus a vencer as suas resistências à biotecnologia.

Por todos esses motivos, a liberação dos transgênicos, sob o controle da agência federal de biossegurança, a CTNBio, é o que se deve esperar do governo Lula.

A demora na definição induz produtores e pesquisadores à ilegalidade

Veículo ESTADO DE S. PAULO		Data 05 JAN 2014		Quadrante	
Página 3-A		Fonte Citada <input checked="" type="checkbox"/> Sem citação		<input type="checkbox"/> Dirigente <input type="checkbox"/> Chefe <input type="checkbox"/> Pesquisador <input type="checkbox"/> Outros empregados	
Composição gráfica <input checked="" type="checkbox"/> Somente texto		<input type="checkbox"/> 02 elementos gráficos <input type="checkbox"/> 03 elementos gráficos <input type="checkbox"/> 04 elementos <input type="checkbox"/> 05 ou mais elementos		Presença do nome <input type="checkbox"/> Capa <input type="checkbox"/> Mês/ano <input type="checkbox"/> Tirulo	
Gênero <input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Carta ao Leitor		<input type="checkbox"/> Nota Informativa <input type="checkbox"/> Nota Operativa		<input type="checkbox"/> Notícia <input type="checkbox"/> Reportagem <input type="checkbox"/> Destaque no Texto <input type="checkbox"/> Rodapé/Legenda	



A hora da inovação, segundo Palocci

Chegou a hora de pôr em primeiro plano a política industrial e tecnológica, isto é, a estratégia necessária ao crescimento de longo prazo. Expansão duradoura não se concretiza apenas com investimento em atividades já conhecidas e já convertidas em rotina. A inovação é componente essencial do desenvolvimento econômico e do crescimento, especialmente quando a produção e o comércio se realizam num ambiente duramente competitivo. Sem abandonar a disciplina fiscal e o cuidado com os fundamentos da economia, é preciso abrir uma nova fase da política. Ninguém o disse tão claramente quanto o ministro da Fazenda, Antônio Palocci, em artigo escrito em colaboração com seu assessor Edmundo Oliveira e publicado domingo no *Estado*.

Mais do que uma nova política, é uma nova cultura que se propõe nesse artigo. "No caso fiscal, já está virando um bem comum a evidência de que o Brasil precisa ser um país organizado, com as contas públicas arrumadas e com uma relação dívida/PIB cadente", afirmam os autores. "No caso da Política Industrial e Tecnológica, ainda vamos levar um tempo

para construir um novo consenso na sociedade brasileira", acrescentam.

A inovação, como preocupação constante e atividade normal, portanto, não deve caber apenas ao governo. É tarefa para o setor público e para o setor privado, que em relação a vários objetivos podem e devem trabalhar em parceria. Não será preciso reinventar as políticas ou descobrir caminhos inteiramente novos. A experiência de vários países mostra fórmulas diversas, e bem-sucedidas, para a condução da pesquisa aplicável ao setor produtivo. Também há exemplos, no País, de iniciativas bem concebidas e desenvolvidas com realismo. Há muito que aprender com as experiências da Embrapa, da Embraer e da Petrobrás, mencionadas no artigo.

O texto contém uma das melhores explicações – e uma das melhores defesas – das "Diretrizes de Política Industrial, Tecnológica e de Comércio Exterior" recentemente apresentadas pelo ministro do Desenvolvimento, Luiz Furlan. Essas diretrizes, segundo o artigo, têm as vantagens da flexibilidade, do foco em setores que transmitem inovações a toda a economia e da abertu-

ra tanto para a pesquisa de ponta quanto para a difusão de tecnologias.

Mas é preciso dar um sentido prático a essas diretrizes. Haverá espaço para debate – necessário, segundo os autores do artigo –, mas também é imprescindível que o governo se organize para atuar em várias frentes. Não basta saber que as verbas públicas virão de fundos setoriais. Serão necessários, como está assinalado no artigo, ajustes de tipo institucional. As entidades financeiras do governo, especialmente o BNDES, terão de se preparar para a aplicação da política. Os autores do artigo não mencionam o Instituto Nacional da Propriedade Industrial (Inpi), mas nenhum trabalho de inovação irá muito longe, sem um eficiente sistema de análise e de registro de patentes. O Inpi foi um dos serviços mais afetados pelo desmonte do setor público, em anos recentes.

O artigo expõe não apenas o que pode ser uma política eficiente de inovação tecnológica, mas também a importân-

cia da articulação do governo em torno de seus objetivos. Não é trabalho para um ou dois ministérios.

Nem todos os ministros estão preparados para participar desse trabalho. O de Ciência e Tecnologia esteve perdido, até agora, assim como o da Educação. Essa falha não será corrigida com a reforma ministerial, se a mudança obedecer a critérios

de acordos partidários e de amizade pessoal.

O presidente, segundo se informou nos últimos dias, considera a hipótese de transferir para o Ministério de Ciência e

A política tecnológica requer mais do que bons amigos no Ministério

Tecnologia, na reforma, o ministro das Comunicações, Miro Teixeira, seu amigo. Substituirá um ministro que apenas enriqueceu o folclore da administração federal por um colega que teve até agora um desempenho lamentável. Se quiser conduzir o País a uma nova etapa de crescimento, o presidente deverá rever alguns de seus critérios e repensar a tarefa de coordenação. Seu ministro da Fazenda oferece, nesse artigo, algumas pistas importantes do que é preciso fazer.

Veículo		Data		Quadrante							
ESTADO DE S. PAULO		19 JAN 2001		<table border="1"> <tr> <td>10</td> <td>18</td> </tr> <tr> <td></td> <td>C</td> </tr> <tr> <td>D</td> <td>E</td> </tr> </table>		10	18		C	D	E
10	18										
	C										
D	E										
Página 3-A		Fonte Citada		<input type="checkbox"/> Dirigente <input type="checkbox"/> Pesquisador <input type="checkbox"/> Sem citação <input type="checkbox"/> Chefe <input type="checkbox"/> Outros empregados							
Composição gráfica		Presença do nome									
<input type="checkbox"/> 02 elementos gráficos <input checked="" type="checkbox"/> Somente texto <input type="checkbox"/> 03 elementos gráficos <input type="checkbox"/> 04 elementos <input type="checkbox"/> 05 ou mais elementos		<input type="checkbox"/> Cabe <input type="checkbox"/> Manchete <input checked="" type="checkbox"/> Título <input type="checkbox"/> Citação <input type="checkbox"/> Destaque no Texto <input type="checkbox"/> Rodapé/Legenda									
Gênero											
<input type="checkbox"/> Crônica <input checked="" type="checkbox"/> Editorial <input type="checkbox"/> Entrevista <input type="checkbox"/> Carta ao Leitor		<input type="checkbox"/> Nota Informativa <input type="checkbox"/> Nota Opinativa		<input type="checkbox"/> Notícia <input type="checkbox"/> Reportagem							

A 'Embrapa industrial'

O grupo de trabalho interministerial que planeja uma política industrial para o País descobriu o ovo de Colombo: por que não usar o modelo que foi utilizado para a criação e desenvolvimento da Embrapa, para criar um organismo que desenvolva e dissemine inovações tecnológicas e difunda nas indústrias o salutar hábito de investir em pesquisa e desenvolvimento?

A idéia foi debatida e aprovada, na quarta-feira passada, pelos ministros que compõem a Câmara de Política Econômica. Os técnicos do governo devem, agora, detalhar as bases conceituais da futura empresa e determinar o seu formato definitivo. Como informou a jornalista Suely Caldas, na edição de sexta-feira do Estado, o presidente Luiz Inácio Lula da Silva e os ministros Luiz Fernando Furlan, do Desenvolvimento, e Antônio Palocci, da Fazenda, apóiam entusiasticamente a proposta.

Numa primeira etapa, será constituída uma rede de integração de vários centros de excelência em pesquisa industrial. Essa rede será articulada com o setor privado e só então esses centros serão unificados em uma única empresa. Foi assim que, há 31 anos, foi criada a Embrapa.

A constituição e integração da rede de institutos de pesquisa é necessária, como salientou o coordenador do grupo de política industrial e presidente do Ipea, Glaucio Arbix, porque o governo não dispõe de instrumentos que dêem racionalidade à ação dos centros de pesquisa na área industrial que, dispersos e sem atuação coordenada, hoje não produzem a sinergia necessária para que o País vença o fosso tecnológico que o separa de países com grau de industrialização semelhante à nossa.

Arbix adverte que a futura empresa não será um cabide de empregos, com ramificações por todos os Estados. Será, isso sim, um organismo voltado pragmaticamente para a redução da dependência tecnológica do País, fomentando a criação e uso de novos processos industriais pelo setor privado.

De fato, a idéia da criação de uma "Embrapa industrial" é boa, mas a sua implementação demandará cuidados para que não se torne um instrumento de novas e perniciosas re-

servas de mercado, nem se limite aos estreitos horizontes da substituição de tecnologias, política que pode produzir bons resultados de curto prazo para as contas nacionais, mas em prazos maiores pouco contribui para a inovação industrial.

O grupo de técnicos que traçará os contornos da nova empresa deverá, também, levar em consideração as diferenças que existem entre as atividades agropecuária e industrial, no que concerne à disseminação de tecnologia. As empresas que desenvolvem técnicas agrícolas - sementes e fertilizantes, por exemplo - dependem basicamente da propagação de seus métodos, para prosperar. Já as inovações industriais têm no sigilo e no domínio exclusivo um fator multiplicador de lucros. Em vista dessas peculiaridades, pretende-se que a "Embrapa industrial" se concentre basicamente no desenvolvimento de técnicas intermediárias de produção, essas sim com boa capacidade de difusão entre as indústrias. Além disso, essas técnicas intermediárias podem funcionar como incentivo para que as indústrias invistam em tecnologia de ponta.

Hoje, as indústrias nacionais preferem comprar a tecnologia disponível no mercado - que quase nunca é a mais avançada. O desenvolvimento de tecnologia autóctone é uma aventura cara e o País não se preparou adequadamente para essa etapa crucial da industrialização. Não há incentivos nem finan-

ciamentos para que o setor privado invista em pesquisa e desenvolvimento e uma legislação complicada e uma burocracia enrolada levam ao desânimo as

Uma boa idéia, mas que se deve implantar com objetivos bem definidos

empresas que se dispõem a desenvolver tecnologias próprias. Em consequência, o pouco de tecnologia que se produz no Brasil é resultado do esforço de empresas estatais e de universidades públicas. É esse quadro pouco favorável ao crescimento do País que a futura empresa pode mudar, assim como a Embrapa transformou o agronegócio brasileiro.

II. Les cinq editoriaux de l' "O Globo"

Veículo		Data: 7 MAR 2001		Quadrante	
Página: 6-OPINIÃO		Fonte Citada: <input checked="" type="checkbox"/> Sem citação		<input type="checkbox"/> Dirigente <input type="checkbox"/> Chefe <input type="checkbox"/> Pesquisador <input type="checkbox"/> Outros empregados	
Composição gráfica: <input checked="" type="checkbox"/> Somente texto		<input type="checkbox"/> 02 elementos gráficos <input type="checkbox"/> 03 elementos gráficos <input type="checkbox"/> 04 elementos <input type="checkbox"/> 05 ou mais elementos		Presença do nome: <input type="checkbox"/> Capa <input type="checkbox"/> Citação	
Gênero: <input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Nota Informativa <input type="checkbox"/> Artigo <input checked="" type="checkbox"/> Editorial <input type="checkbox"/> Carta ao Leitor <input type="checkbox"/> Nota Opinião <input type="checkbox"/> Reportagem		<input type="checkbox"/> Manchete <input checked="" type="checkbox"/> Destaque no Texto <input type="checkbox"/> Título <input type="checkbox"/> Rodapé/Legenda		<input type="checkbox"/> B <input type="checkbox"/> C <input type="checkbox"/> D <input type="checkbox"/> E	

OPINIÃO

Cerco ideológico

Não é de hoje que a agricultura tem funcionado como o mais eficiente motor da economia brasileira. Ficou distante no passado aquele setor relegado a segundo plano, sinônimo de atraso, preocupado apenas em colher alimentos e despachá-los *in natura* para o consumo. As fazendas tornaram-se pólos de agronegócios, um sistema produtivo cuja cadeia beneficia, e muito, as cidades.

Hoje, 37% dos empregos na economia e 27% do PIB são gerados por essa cadeia. O superávit comercial da agropecuária mantém o país solvente diante dos credores externos — e continua a melhorar os indicadores externos da economia. Deve-se colher este ano mais uma safra recorde: 130 milhões de toneladas, 8% acima dos pouco mais de 120 milhões de toneladas do período anterior. Tudo em meio a enormes ganhos de produtividade: desde 1995, a produção mais que dobrou (125%), enquanto a área plantada expandiu-se apenas 14%.

No centro dessa revolução está a Empresa Brasileira de Pesquisa Agropecuária (Embrapa), cujos projetos têm sido a mola propul-

sora dessa histórica explosão do agronegócio no Brasil. Pois, apesar disso, a Embrapa corre sérios riscos por causa da interferência política no seu trabalho, como noticiado no domingo pelo GLOBO. Grande usina geradora de soluções tecnológicas que passaram a situar o Brasil, segundo estudos da ONU, como o provável maior produtor de alimentos do mundo dentro de 12 anos, à fren-

te dos Estados Unidos, a Embrapa já reduziu o ritmo das pesquisas, por causa da interferência de militantes sindicais. Um dado: em 2002, a empresa registrou 39 patentes; no ano passado, apenas 13.

A História ensina que ideologia não rima com ciência. E sempre que as duas se misturam, a ciência é derrotada. Tudo indica, é o que começa a acontecer na Embrapa. O assunto merece, no mínimo, uma reflexão por parte dos altos escalões do governo. Essa ameaça que paira sobre o futuro da agropecuária é maior que a de qualquer praga. O cerco ideológico à Embrapa, pela sua gravidade, também deve merecer atenção do Congresso e da comunidade científica. Estão em jogo questões estratégicas para o país.

A Embrapa enfrenta a praga da ingerência política

Veículo	O GLOBO	Data	03 FEV 2004	Quadrante	
Página	06 Opinião	Fonte Citada	<input checked="" type="checkbox"/> Sem citação <input type="checkbox"/> Dirigente <input type="checkbox"/> Pesquisador		
Composição gráfica	<input type="checkbox"/> 02 elementos gráficos <input type="checkbox"/> 03 elementos gráficos <input type="checkbox"/> 04 elementos <input type="checkbox"/> 05 ou mais elementos		<input type="checkbox"/> Chefe <input type="checkbox"/> Outros empregados		
Gênero	<input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Nota Informativa <input type="checkbox"/> Nota <input type="checkbox"/> Artigo <input checked="" type="checkbox"/> Editorial <input type="checkbox"/> Carta ao Leitor <input type="checkbox"/> Nota Opinião <input type="checkbox"/> Reportagem		Presença do nome	<input type="checkbox"/> Cabe <input checked="" type="checkbox"/> Citação <input type="checkbox"/> Manchete <input type="checkbox"/> Destaque no Texto <input type="checkbox"/> Título <input type="checkbox"/> Rodapé/Legenda	

OPINIAO

Na contramão

A surpresa causada pelo anúncio de que o governo federal prepara-se para fazer uma gigantesca contratação de funcionários públicos é proporcional ao equívoco da iniciativa. E o erro não está apenas na magnitude dos números envolvidos; ele também transparece em conceitos frágeis que tentam justificar essa espécie de megatrem da alegria.

A operação faz parte de um plano pensado e articulado há algum tempo. Pois, das 65 mil contratações, quase 25 mil já haviam sido autorizadas no ano passado. Para se ter uma idéia do tamanho do gulchê de recrutamento que o Palácio do Planalto está abrindo, basta lembrar que nos oito anos do governo de Fernando Henrique Cardoso foram empregados 52 mil novos servidores.

Deve ser considerado, por óbvio, que a máquina pública precisa de reposição em setores-chave. Não faz sentido debilitar as estruturas de atendimento à população. Não se justifica, por exemplo, deixar de repor pessoal nas polícias ou nas áreas de saúde. Seja em que nível for: federal, estadual ou municipal. A discussão é outra. E querer travá-la em torno de conceitos como "estado forte versus estado mínimo" é contaminar de ideologia um tema que merece reflexão serena e fundamentada. Também é tomar atalhos sem direção fazer comparações de índices de proporção de funcionários públicos em relação à população total entre o Brasil e potências globais ou com outros países que ostentam uma saúde fiscal em nada semelhante à brasileira.

Em vez de contratar o governo deveria reformar

O importante não é o estado ser forte ou mínimo. O importante é um estado eficiente, e presente onde necessário: no atendimento à população. Hoje, há 792 mil servidores federais, 46% menos que o 1,48 milhão de 1989. O custo, porém, para a

União continua elevado. Afinal, boa parte dos funcionários que saíram da folha de pagamento entrou nas contas deficitárias da Previdência.

Assim, em vez de fazer supercontratações, o governo deveria retomar a reforma administrativa que vinha sendo executada, para aumentar a eficiência da máquina. Além disso, aconselha-se ao Planalto esclarecer todas as dúvidas em torno do assunto. Pode começar por atender aos pedidos de informação do PFL e do PSDB.

Ciência livre

O substitutivo do ex-líder do governo, Aldo Rebelo, agora ministro da Articulação Política, ao projeto que cria a política de biossegurança do país tem recebido de setores do próprio governo críticas tanto acerbas quanto indevidas. A questão central é que o projeto original enviado ao Congresso previa que a palavra final sobre pesquisas com organismos transgênicos caberia aos ministérios da Saúde, do Meio Ambiente e da Agricultura; e o substitutivo transfere o poder de decisão à Comissão Técnica Nacional de Biossegurança (CTNBio).

É isso que o ministro Miguel Rossetto, do Desenvolvimento Agrário, e Marina Silva, do Meio Ambiente, rejeitam, por julgar que deve caber aos ministérios a competência não só para liberar ou não a comercialização de organismos geneticamente modificados, mas também de autorizar pesquisa científica.

É difícil entender por que se poderia imaginar que uma questão eminentemente técnica deva ficar subordinada a políticos. A explicação talvez esteja numa visão distorcida, ainda que política-

mente correta e bastante difundida, que considera os transgênicos inteiramente nocivos ao meio ambiente, à saúde e a tudo o mais, e se recusa a reconhecer neles quaisquer aspectos positivos.

Cercear o plantio e a comercialização de transgênicos já é por em risco o avanço notável alcançado nos últimos anos pela agricultura brasileira, principalmente no que se refere à soja. Criar empecilhos à atividade dos cientistas, quando em grande parte esse progresso foi resultado justamente da pesquisa científica — veja-se o trabalho desenvolvido pela Embrapa — seria um duro golpe à capacidade dos produtos agrícolas nacionais de manterem sua competitividade nos mercados internacionais.

A liberação e a fiscalização do plantio de culturas transgênicas têm merecido uma discussão complexa; mas no que se refere à pesquisa, não há o que discutir. A imposição de controles externos à busca do conhecimento é sempre inadmissível, e deve ser classificada de destrutiva quando esse trabalho tem se mostrado economicamente tão positivo para o país.

A pesquisa de transgênicos não pode ficar subordinada aos políticos

Veículo		GLOBO		Data 25 DEZ 2003		
Página 6		Fonte Citada		<input type="checkbox"/> Dirigente <input type="checkbox"/> Pesquisador <input checked="" type="checkbox"/> Sem citação <input type="checkbox"/> Chefe <input type="checkbox"/> Outros empregados		
Composição gráfica		<input type="checkbox"/> 02 elementos gráficos <input type="checkbox"/> 04 elementos <input checked="" type="checkbox"/> Semente texto <input type="checkbox"/> 03 elementos gráficos <input type="checkbox"/> 05 ou mais elementos		Presença do nome		
Gênero		<input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Nota Informativa <input type="checkbox"/> Notícia <input type="checkbox"/> Artigo <input checked="" type="checkbox"/> Editorial <input type="checkbox"/> Carta ao Leitor <input type="checkbox"/> Nota Operativa <input type="checkbox"/> Reportagem		<input type="checkbox"/> Capa <input type="checkbox"/> Citação <input type="checkbox"/> Manchete <input checked="" type="checkbox"/> Destaque no Texto <input type="checkbox"/> Título <input type="checkbox"/> Rodapé/Legenda		

TEMA EM DISCUSSÃO: *Transgênicos*

NOSSA OPINIAO

Ciência a salvo

Permanece no projeto do governo sobre transgênicos, que está tramitando na Câmara dos Deputados, a criação do Conselho Nacional de Biossegurança, que tomará as decisões finais sobre a liberação de organismos geneticamente modificados. Será um conselho integrado por ministros, portanto um órgão político, ao qual a Comissão Técnica Nacional de Biossegurança (CTNBio) apenas prestará assessoria. É um claro equívoco, porque o tema é complexo e as decisões a seu respeito exigem conhecimento técnico especializado.

No entanto, o prejuízo é parcial e poderá ser sanado mais adiante, porque o assessor do projeto, deputado Aldo Rebelo, agindo com sabedoria, cuidou de separar plantio e comercialização de transgênicos da autorização para pesquisas em biotecnologia, que

continuará a cargo da CTNBio.

Isso permitirá que a ciência e a tecnologia — que, particularmente graças ao trabalho da Embrapa, já deram contribuições muito valiosas para a agricultura brasileira, especialmente no caso da soja — continuem progredindo, livres dos obstáculos impostos pela equivocada visão ideológica da questão que continua a prevalecer em parte do governo.

Desta forma, quando a racionalidade voltar a se impor (o que certamente não tardará muito, e provavelmente tomará a forma de uma imperiosa necessidade de preservar mercados de exportação ameaçados pela competição de safras transgênicas de outros países), será possível recuperar o atraso sem maiores dificuldades. Não ficaremos dependentes da importação de tecnologia desenvolvida no exterior por grandes empresas multinacionais.

Veículo GLOBO		Data 11 ABR 2004	Quadrante <table border="1"> <tr> <td>A</td> <td>B</td> </tr> <tr> <td>X</td> <td>C</td> </tr> <tr> <td>D</td> <td>E</td> </tr> </table>	A	B	X	C	D	E
A	B								
X	C								
D	E								
Página 06	Fonte Citada <input checked="" type="checkbox"/> Sem citação	<input type="checkbox"/> Dirigente <input type="checkbox"/> Chefe <input type="checkbox"/> Pesquisador <input type="checkbox"/> Outros empregados							
Composição gráfica <input checked="" type="checkbox"/> Somente texto	<input type="checkbox"/> 02 elementos gráficos <input type="checkbox"/> 03 elementos gráficos <input type="checkbox"/> 04 elementos <input type="checkbox"/> 05 ou mais elementos	Presença do nome <input type="checkbox"/> Capa <input checked="" type="checkbox"/> Citação <input type="checkbox"/> Manchete <input type="checkbox"/> Destaque no texto <input type="checkbox"/> Título <input type="checkbox"/> Rodapé/Legenda							
Gênero <input type="checkbox"/> Artigo <input checked="" type="checkbox"/> Editorial	<input type="checkbox"/> Crônica <input type="checkbox"/> Entrevista <input type="checkbox"/> Carta ao Leitor	<input type="checkbox"/> Nota Informativa <input type="checkbox"/> Nota Opinativa <input type="checkbox"/> Notícia <input type="checkbox"/> Reportagem							

Dogmas

Uma tendência que poderíamos chamar de fundamentalismo ambientalista passou a constituir um grave empecilho à pesquisa científica, à continuidade da notável trajetória de sucesso da agricultura brasileira e ao próprio desenvolvimento econômico do país.

A premissa a partir da qual se estrutura esse pensamento é perfeitamente lógica: o meio ambiente está sob risco no Brasil, e grandes obras (como a construção de usinas hidrelétricas e estradas), assim como novas tecnologias, como a modificação genética de determinadas lavouras, podem ter efeitos danosos que é preciso evitar.

Mas esse tipo de raciocínio deveria conduzir à adoção da maior cautela e de exames criteriosos para o licenciamento ambiental de projetos, assim como de estímulo à pesquisa científica para avaliação precisa dos efeitos que podem ter produtos novos como os transgênicos. Em lugar disso, o que se registra é um esforço para impedir o avanço da ciência e da tecnologia nacionais e a criação de toda sorte de obstáculos burocráticos à concessão de licenças

para projetos de engenharia de maior importância.

Na agricultura, a ameaça criada é dupla: de um lado tenta-se evitar, por meio de uma legislação restritiva, que o Brasil acompanhe a evolução tecnológica que avança a passos largos em outros países; enquanto o órgão que até aqui foi o principal motor do progresso brasileiro nessa

área, a Empresa Brasileira de Pesquisa Agropecuária (Embrapa), passou a sofrer, desde o ano passado, tamanho cerceamento em sua pesquisa, devido à interferência ideológica, que já reduziu drasticamente seu ritmo de trabalho científico.

Ninguém pode defender o desenvolvimento a todo custo, com o mais completo descaso pelas questões ambientais. Mas é preciso buscar um sensato meio-termo e, sobretudo, evitar o tratamento ideológico de temas que são de natureza técnica e científica. Caso contrário, essa visão tende a resvalar para um dogmatismo extremamente danoso. Hoje, na realidade, isso já deixou de ser uma ameaça para se tornar um obstáculo real, que o país precisa derubrar para não perder o rumo do progresso.

Defender o
meio ambiente
é questão de
sensatez, não
de Ideologia