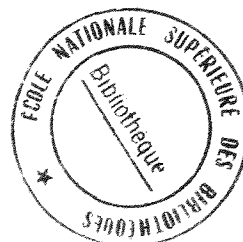


**Ecole Nationale
Supérieure de
Bibliothécaires**

**Diplôme Supérieur
de Bibliothécaire**

**Université
Claude Bernard
Lyon I**

**DESS Informatique
Documentaire**



Note de synthèse

*Evaluation des ressources documentaires: il s'agit de faire une synthèse des études réalisées et des méthodes utilisées pour l'évaluation des banques de données documentaires. On pourra se limiter à l'un des aspects suivants: * comparaison et évaluation des performances des logiciels d'interrogation: qualité des logiciels. * Evaluation et comparaison du contenu des bases méthodes et résultats.*

Christine POISSON

*Sous la direction de Christine ANDRE et
Danièle ROGER, ENSB.*

1991

Evaluation des ressources documentaires: il s'agit de faire une synthèse des études réalisées et des méthodes utilisées pour l'évaluation des banques de données documentaires. On pourra se limiter à l'un des aspects suivants: * comparaison et évaluation des performances des logiciels. * Evaluation et comparaisons du contenu des bases: méthodes et résultats.

Christine Poisson

RESUME: L'évaluation des bases de données documentaires peut s'effectuer à deux niveaux: le support et le contenu. Des méthodes générales et spécifiques se sont développées. Une synthèse en est ici proposée, elle comporte aussi une partie bibliographique.

DESCRIPTEURS: Base de données; étude comparative; évaluation; recherche documentaire; évaluation performance.

ABSTRACT: The evaluation of online retrieval databases could be done at two levels: information storage medium and content. General and specific methods have been developed. A synthesis is proposed here, it contains a bibliographic part too.

KEYWORDS: Database; comparative study; evaluation; retrieval information; performance evaluation.

INTRODUCTION

L'évaluation des bases de données est devenue au fil du temps un outil de gestion pour les producteurs et les serveurs mais aussi un moyen pour l'utilisateur qui entreprend une recherche d'effectuer le ou les bons choix. La prolifération des bases de données (environ 4000 à l'heure actuelle) incite les spécialistes à établir des critères de comparaison.

L'objectif de Mesdames André et Roger en proposant ce sujet était de faire le point sur la situation. La question est de savoir quelles sont les méthodes de comparaisons et d'évaluations couramment employées actuellement. La connaissance de ces nouvelles techniques pourraient être éventuellement intégrée au cours de recherche documentaire automatisée.

Ce sujet se limitera à la recherche bibliographique de ces dernières années. Il a été décidé avec Madame André de réduire le sujet à un seul domaine. J'ai choisi le secteur médical. Dans le but de vérifier si les méthodes employées sont partout identiques, j'ai prolongé l'étude en intégrant des références d'autres domaines.

La méthodologie de recherche automatisée et manuelle forme une première partie. La synthèse se divise, elle, en deux sous-ensembles: l'évaluation des supports des bases de données (les serveurs et les nouveaux supports: le CD Rom) et l'évaluation du contenu des bases. Il s'agit là de faire le point sur les théories générales, sur les comparaisons elles-mêmes et sur des comparaisons plus spécifiques (qualité, indexation, et couverture). Les éléments qui ont contribué à réaliser cette recherche sont rassemblés dans la bibliographie.

I- METHODOLOGIE DE RECHERCHE.

Cette recherche s'est déroulée en deux phases qui se sont entre-croisées: une recherche automatisée et une autre manuelle.

A- La recherche automatisée.

Cette recherche automatisée a tout d'abord été une étape de mise en route du sujet. Une interrogation en ligne a été effectuée dans la base scientifique Pascal sur le serveur Questel, d'autres recherches ont ensuite complété la première: de nouveau sur Pascal (en ligne et sur CD Rom) et dans Lisa (library information science abstracts).

1- La stratégie.

La stratégie de la recherche en ligne sur Pascal se décompose en quatre étapes. Cette recherche a été effectuée en français comme suit:

- Q1: Base donnée ou Banque donnée
- Q2: 1 et évaluation/FD
- Q3: 2 et (science 2AV information)
- Q4: 3 et étude comparative

22 références ont ainsi été obtenues.

Compte tenu des références trouvées parallèlement au cours de la recherche manuelle ,les résultats de la première interrogation n'apparaissaient pas comme totalement satisfaisants. Il a donc été décidé d'entreprendre une nouvelle série de recherches en ligne dans Pascal afin de compléter la bibliographie.

À partir de références pertinentes déjà obtenues, Mme Roger a effectué une recherche pour retrouver des mots-clés importants.

Celle ci a débouché sur une nouvelle interrogation.Elle a été effectuée sur le CD Rom Pascal qui couvre les quatre dernières années de la base en ligne. J'ai employé les termes anglais pour une recherche dans les champs titre et descripteurs:

- Q: (indexing ou coverage ou evaluation ou comparison ou information retrieval) et database.

Une quinzaine de références supplémentaires ont ainsi été obtenues.

La recherche dans Lisa prévue au départ en ligne a été effectuée manuellement. Les deux dernières années de la forme papier ont été systématiquement dépouillées ainsi que les deux premiers fascicules de cette année. Cette recherche a été menée grâce à une série de cinq descripteurs anglais:

- * Evaluation
- * Online information retrieval
- * Retrieval performances measures
- * Comparative studies
- * Databases

Cette recherche a débouché sur une douzaine de nouvelles références.

2- Les résultats et commentaires.

Ces différentes recherches ont donc donné 49 références bibliographiques. Or elles ne sont pas toutes très pertinentes.

La première recherche effectuée dans Pascal a donné de bons résultats au niveau du contenu mais un problème se pose au niveau des dates de parution des différents articles obtenus. 16 d'entre eux sur 22 sont antérieurs à 1985, ce qui semble trop vieux du point de vue de l'évolution de la technologie, des serveurs, des logiciels et du contenu des bases de données. Toutefois, on peut considérer deux cas:

* Des références qui sont vieillissantes lorsqu'il s'agit de comparer deux bases de données entre elles au niveau du contenu ou des serveurs sur lesquelles elles sont implantées.

* Mais des références qui demeurent plus pertinentes lorsqu'elles concernent les techniques d'évaluation qui ont peu varié.

La seconde phase d'interrogation de la base Pascal à partir des mots clés a également donné des résultats mitigés. Pertinents au niveau des dates, la moitié d'entre eux ne le sont pas quant au contenu. Mais cette recherche a été intéressante en plus des nouvelles références obtenues dans la mesure où elle a permis de déterminer un certain nombre de mots-clés repris pour la recherche dans Lisa.

Cette dernière est tout à fait pertinente: le dépouillement manuel de la forme papier se traduisant par le choix direct des références intéressantes.

D'une façon générale, les résultats obtenus par la recherche en ligne sont donc mitigés. Ils ont permis d'orienter le sujet de trouver des références intéressantes mais dont le défaut majeur réside dans leur relative ancienneté.

B- La recherche manuelle.

1- Orientation de la recherche

Cette recherche se décompose en deux sous ensembles: les monographies et les périodiques.

* Les monographies

Il n'existe pas de monographies sur l'évaluation des bases proprement dites. Il existe deux types d'ouvrages:

- Les monographies sur les bases de données en ligne et les répertoires de bases

- Les ouvrages généraux sur l'évaluation de la recherche documentaire.

J'ai donc feuilleté ce type de livres pour déterminer s'il pouvait être intéressant d'inclure certains d'entre eux dans mes recherches.

* Les périodiques

Pour ce type de sujet, ce sont les articles de périodiques qui sont les plus pertinents de par leur fraîcheur d'information. Il existe un certain nombre de périodiques spécialisés dans la recherche documentaire et dans les bases de données. On peut citer en exemple Database, Online, Online Review, journal of the American society for information science. Ces journaux ont été dépouillés de façon systématique pour ces trois dernières années.

2- Les résultats et les commentaires

Ces recherches ont permis de trouver quelques ouvrages généraux qui dans un chapitre donnaient quelques points de méthodologie pour l'évaluation des bases de données. Ces méthodes datent des années soixante-dix, elles ont parfois été complétées mais les principes de base restent toujours identiques tout au moins dans ces livres.

La vitesse d'évolution rendrait pratiquement un ouvrage obsolète dès sa sortie. On trouve en conséquence davantage d'articles de périodiques sur le sujet. Le dépouillement systématique a donné de bons résultats. Les références d'abord cherchées dans le domaine médical ont été partiellement complétées par d'autres domaines pour vérifier si les méthodes employées dans un secteur de l'activité documentaire restait valable pour les autres.

Ces références obtenues directement sont ensuite une nouvelle source d'investigations grâce aux références bibliographiques elles mêmes mentionnées à la fin de ces articles .

L'ensemble de cette recherche bibliographique a donc permis de retrouver 66 références concernant l'évaluation des bases de données. Davantage orientée vers le domaine médical, la documentation rassemblée sur la comparaison des bases dans d'autres domaines n'est donc que complémentaire.

Cette recherche est inégale quant au contenu des documents obtenus. Certains aspects demeurent plus développés que d'autres. Les comparaisons de bases de données entre elles sont très nombreuses . Les évaluations des CD Rom ne sont pas non plus négligées: le développement rapide de ce nouvel outil documentaire et sa prolifération sur le marché explique le nombre des études réalisées récemment. Les évaluations spécialisées sur un aspect précis des bases comme par exemple l'indexation moins importantes en nombre sont beaucoup plus techniques et font apparaître de vrais spécialistes de la question.

II- LA SYNTHÈSE.

A- Les supports des bases de données.

Lorsqu'une base apparaît sur un nouveau serveur, il y a une question que le spécialiste est en droit de se poser. Quelle est la différence avec la même base sur un autre serveur ou support? Qu'apporte-t-elle de plus? Des méthodes d'évaluation sont mises en place afin de répondre à ce type de questions.

1- L'évaluation des serveurs.

La méthode de base employée est toujours la même: une ou plusieurs recherches sont effectuées sur les différents serveurs. Chacun d'entre eux possède ses propres caractéristiques, les résultats obtenus sont alors étudiés tant au niveau du contenu que des coûts. Les discussions s'organisent autour des différences entre les systèmes. Certains critères sont automatiquement pris en ligne de compte pour tout type d'évaluation: ce sont les critères de coût et des critères dit généraux.

En matière de coût, plusieurs éléments sont importants: le prix de la communication, le coût de visualisation des documents. La rapidité du système peut alors se révéler non négligeable.

Parmi les critères généraux, on peut citer la date de départ de la couverture documentaire (pour les recherches rétrospectives), l'intervalle entre deux mises à jour, les possibilités offertes par le système de créer des profils, ou encore l'aide en ligne.

Les discussions vont permettre de distinguer des différences mineures et majeures (4).

Les premières se situent au niveau des caractéristiques propres de la base:

- * Des variations dans la terminologie. Dans certaines bases ces différences sont réduites dans la mesure où l'emploi d'un langage contrôlé est pratiquement obligatoire. On peut citer le MESH (medical subject headings) pour Medline.

- * Des formats de sorties des documents propres.

- * Le passage en revue de tous les champs fait également apparaître de petites différences quant au contenu. Le champ titre peut par exemple ne contenir les titres de périodiques qu'en abrégé.

La capacité de recherche englobe elle, les différences majeures (2,3):

- * Les champs interrogeables forment un premier sous ensemble. La recherche dans le champ des descripteurs, dans d'autres champs tels que le titre, l'année de parution, ou la recherche en texte libre dans tous les champs caractérisent les serveurs

* La possibilité d'explorer les termes lorsque la structure du thésaurus est arborescente, et la possibilité de limiter le sujet à un domaine précis (comme par exemple humain et animal dans Medline) permet aussi de distinguer les serveurs.

* Les limites de la logique booléenne sont un autre élément de distinction.

Les responsables des évaluations donnent en guise de conclusion, quelques éléments pour le choix d'un serveur mais ils restent très généraux. Selon les moyens dont l'établissement dispose, l'un d'entre eux peut être préféré tandis que pour des recherches très pointues un autre serveur peut s'imposer. Mais la concurrence tend à réduire les différences majeures entre les systèmes: les gros défauts disparaissent et chacun met en avant ces propres atouts. L'habitude d'interrogation d'un serveur reste aussi non négligeable lorsqu'il s'agit d'effectuer des recherches.

2- Les nouveaux supports: les CD Rom: comparaison avec les bases en lignes.

A l'heure où les CD Rom se développent de façon intensive et où les bibliothèques et autres centres de documentation s'équipent, le choix n'est pas toujours facile. Des études comparatives sont réalisées régulièrement. Elles comparent des CD Rom entre eux mais aussi aux bases en ligne correspondantes. Ces études peuvent contribuer au choix du produit recherché. Il est vrai que le CD Rom apparait comme beaucoup plus simple d'accès et que la recherche en ligne reste le privilège d'un petit nombre.

On peut distinguer les méthodologies d'évaluation proposées dans les périodiques et les critères mis en avant par les spécialistes qui doivent permettre de faire des comparaisons.

Les évaluations:

Quelques précautions s'imposent avant de pouvoir évaluer et comparer. La base doit ainsi être disponible sur plusieurs serveurs. Les différentes versions doivent être testées sur un matériel ayant la même configuration. Une réelle comparaison quantitative demande en plus une sélection très attentive des termes de recherche, faisant état des différences dans la stratégie d'indexation entre les versions des CD Rom. En tenant compte de ces conditions, même les CD Rom avec des bases différentes peuvent être comparées. Les tests actuels peuvent révéler d'importantes différences (18).

Une fois les conditions posées, les évaluations se déroulent de façon très proches. Les chercheurs sont répartis en paires et interrogent la base sur CD Rom puis la version en ligne lorsque les deux sont comparés. Certaines études incluent également des étudiants dans leurs groupes de recherche (17).

Une ou plusieurs séries de questions déterminées à l'avance leur sont alors fournies. Celles ci sont soit tirées au sort parmi des questions de références, soit posées par un spécialiste du domaine. Ce peut être aussi des questions régulièrement posées par les étudiants.

Les critères de choix:

Ce sont les critères locaux qui feront pencher la balance dans un certain sens (lorsqu'il s'agit de s'équiper) mais il existe quelques éléments généraux qui ont été mis en avant. Je proposerai quatre catégories:

* Les critères de matériels et de coûts. Est-il possible d'utiliser plusieurs CD Rom sur le même matériel? Le matériel qu'un établissement possède déjà est-il bien adapté? Le CD Rom est-il facile à installer? Au niveau du coût, il faut tenir compte du coût de souscription et du nombre d'utilisateurs potentiels (la recherche est en ligne ne serait-elle pas moins couteuse si le potentiel est faible?) (17).

* Les critères de volume et de couverture (18). Il est vrai que le CD Rom possède une capacité réduite par rapport à la base en ligne. Mais les différences sur un même support ne sont pas à négliger. Le volume de la base (un ou plusieurs disques), les dates de couverture peuvent révéler de grandes différences. La restriction de la base à un certain type de périodiques (par exemple réduction de la base aux périodiques de langue anglaise) est un critère d'évaluation. Enfin l'intervalle de mise à jour est un élément de poids pour comparer les performances des CD Rom.

* Les mesures au niveau de la recherche sont capitales pour pouvoir évaluer ces bases de données. On peut citer ceux qui sont systématiquement étudiés (17). La vitesse de recouvrement de l'information est toujours prise en compte, la comparaison avec les bases en ligne se traduit par des variations par rapport à la personne qui interroge. Les utilisateurs réguliers obtiennent de meilleurs résultats en ligne contrairement aux étudiants. Mais d'une façon générale l'accès en ligne est beaucoup plus rapide. Cela fait apparaître un autre élément de réponse: la convivialité du système. Les CD Rom offrent un certain confort pour les utilisateurs temporaires. La recherche par menus déroulants leur en facilite l'accès. Il s'agit alors de les étudier et de les comparer entre eux. Un manuel d'utilisation bien réalisé et une aide en ligne contribuent à cette évaluation. Les champs interrogeables, la méthode d'indexation, l'emploi de la logique booléenne sont indispensables pour qu'une comparaison valable soit effectuée. On peut ajouter les limites qu'on désire imposer au système. Enfin une stratégie peut-elle être sauvée et utilisée directement pour une autre recherche (sur un autre disque de la base) (10,17,18).

* La sortie des données forme une quatrième catégorie. Les différents formats proposés y sont étudiés, la possibilité d'impression, et le téléchargement. Peut-on par

exemple intégrer les références obtenues dans un programme de stockage (16,18).

Certains articles proposent aux lecteurs des questions auxquelles ils se doivent de répondre pour évaluer les différents disques. D'une manière générale les CD Rom vont continuer à évoluer, ce produit jeune se heurte directement aux bases en ligne qui ont elles vingt ans d'expérience. La comparaison doit tenir compte de ce qu'on veut obtenir: une certaine convivialité accessible à tous ou un outil très performant pour une recherche pointue et très récente (on choisira alors la base en ligne).

B- L'évaluation du contenu des bases de données.

1- Les théories générales d'évaluation.

Si les théories dans leurs principes fondamentaux ont peu varié au cours des dernières années, les évaluations sont devenues un véritable outil de gestion pour les producteurs et les serveurs.

Il existe trois grands types d'évaluation (22,23,25):

* L'évaluation de l'efficacité d'une base de données. Les macro-évaluations tendent à répondre à la question: quel est le degré d'efficacité du système. Les résultats renseignent sur le fonctionnement du service mais pas sur la manière de l'améliorer. Les micro-évaluations (les plus pratiquées) permettent de poser un diagnostic analytique. Elles répondent à la question: pourquoi le système fonctionne à son niveau actuel, qu'est ce qui peut l'améliorer?

* L'évaluation coût-efficacité. Elle mesure l'efficacité du système en termes de coûts (efforts, temps, argent). Elle calcule le coût par référence pertinente et le coût par nouvelle référence pertinente.

* L'évaluation coût-avantages. Il s'agit du rapport coût prestation / avantages résultants du bénéfice du service. Cette analyse s'avère ardue.

Une méthodologie en plusieurs étapes est décrite elle aussi de façon très globale: définition d'un champ d'évaluation, conception du programme d'évaluation, exécution, analyse et interprétation, modification du système après les résultats. Ces évaluations doivent prendre en considération non seulement la base mais aussi son accès.

Ces théories générales débouchent sur une sélection de critères et sur des mesures d'efficacité.

Treize grands domaines sont déterminés (21):

- La couverture en est le premier, car si la base ne couvre pas le champ d'intérêt aucun autre critère ne peut être appliqué.

Le type de documents indexés, l'orientation de la base (grand public ou recherche), les dates de couverture, la politique de sélection, la proportion de documents internationaux, et la multidisciplinarité de la base de données en sont les principaux éléments.

- Le deuxième critère concerne la taille: le nombre de références, leur croissance, la division en sous-fichiers.

- La disponibilité locale des documents ne doit pas être négligée.

- La circulation des documents est essentielle pour certains chercheurs. L'intervalle entre deux mises à jour, le temps d'intégration des références et l'exactitude des informations peut déterminer de la valeur d'une base.

L'indexation pèse lourdement sur l'évaluation. La première question est de savoir dans quels champs le vocabulaire contrôlé est employé. Il faut vérifier s'il existe un résumé, sa nature, la profondeur et l'exhaustivité de l'indexation. Des données comme les titres des périodiques ou les noms sont elles standardisées.

- Les pourcentages d'erreurs doivent aussi apparaître dans une évaluation. Ces erreurs peuvent être factuelles, typographiques, ou d'indexation.

- Le traitement des concepts de recherche: certaines bases donnent uniquement des publications de recherche.

- L'enregistrement et la structure des fichiers sont le huitième critère d'évaluation. La recherche par défaut dans certains champs est par exemple classée dans cette catégorie.

- Les possibilités d'impression et de sorties sont aussi étudiées. Le système est-il capable de trier les références? Quel est le tri par défaut? Quels sont les formats d'édition? Ce sont des questions que se posent les évaluateurs.

- Les coûts monétaires sont souvent le premier critère du client mais il faut y ajouter les coûts en temps. Le temps requis pour une interrogation détermine aussi de l'efficacité du système. Or tous ces coûts sont en évolution permanente. Très peu d'études sont réalisées sur ce sujet car elles risqueraient d'être obsolètes avant leurs sorties.

- Les aides disponibles sont importantes pour l'utilisateur du système. Plusieurs outils lui sont proposés: le manuel, le thésaurus, le plan de classification... Il peut aussi disposer d'une aide en ligne mais son coût élevé contribue à le déconseiller.

- La comparaison avec une version sur papier lorsqu'elle existe peut se révéler intéressante. La différence majeure se situe au niveau du nombre de points d'accès.

- Enfin le système propose-t-il des articles de synthèse sur un domaine (review)? Ils sont alors analytiques ou comparatifs.

L'évaluation d'une base de données ce peut être aussi calculer son efficacité. On peut mesurer ou estimer les taux de rappel et de précision qui sont atteints sur un échantillon représentatif. "Ces taux révèlent le degré auquel le système a été capable d'explorer la base de données pour en extraire les références pertinentes et en laissant de coté les autres" (F. Lancaster) (22,23).

Pour obtenir de bons éléments de comparaison, il faut utiliser ces deux ratios.

* Le taux de rappel mesure la proportion de documents pertinents qui ont été capturés par la formulation de la recherche. Il se calcule comme suit: Nombre de documents pertinents trouvés / Nombre de documents pertinents dans la base.

* La précision évalue la pureté de la sortie: l'étendue des documents trouvés qui sont pertinents: Nombre de documents pertinents / Nombre de documents trouvés.

Ces calculs faciles en théorie se heurtent à la façon de mesurer les documents pertinents dans la base. Lancaster propose d'effectuer une recherche équivalente dans plusieurs systèmes qui pourront être sous forme papier. On peut également coupler la recherche d'un spécialiste du domaine et d'un spécialiste de l'information. P. Vigil suggère une technique qui utilise l'opérateur booléen sauf. Grâce à trois recherches, il élimine les éléments redondants. Cette méthode permet aussi de donner les éléments de vocabulaire associés au document (27).

Des études plus récentes comme celle de Mc Carn contribuent à faire évoluer ces calculs en offrant une plus grande précision et en les complétant. Evaluer une base à plusieurs années d'intervalle, comparer sur le même plan les ratios de deux bases qui ne sont pas tout à fait semblables dans leur configuration (24).

Ces critères généraux posés, il ne reste plus qu'aux évaluateurs à les appliquer de façon globale ou séparée.

2- Les comparaisons générales et plus particulièrement quantitatives.

Ce qui était vrai au début de cette synthèse à propos des serveurs l'est aussi pour les comparaisons générales. Lorsqu'une nouvelle base apparaît sur un marché déjà bien pourvu (environ 4000 bases actuellement et près de 80 dans le domaine médical), on peut se demander ce qu'elle propose de plus par rapport à ces concurrentes directes.

Entreprendre une évaluation de grande envergure prend du temps, le travail se compte en mois. On peut donc distinguer deux niveaux d'évaluation: les études en profondeur et les comparaisons plus sommaires qui ne se penchent que sur un point précis.

Pour une bonne exploration, il faut définir les composantes essentielles des bases comparées, identifier les variables affectant la performance et désigner une méthode d'obtention de résultats. L'évaluation se fait en deux temps:

* Une comparaison générale de la base ou le passage en revue des éléments principaux qui sont établis par le producteur. On peut citer par exemple le nombre de références enregistrées, la croissance annuelle de la base, le type de documents proposés... Beaucoup d'évaluations s'arrêtent à ce niveau ou ne vont pas vraiment plus loin (32,36,41).

* Une comparaison des résultats doit être associée. Elle suit alors la méthodologie utilisée pour l'expérimentation (34).

Cette méthodologie repose sur le choix des questions qui seront posées au cours de l'interrogation en ligne. On note deux méthodes. La première consiste à choisir au hasard les questions parmi une liste de questions dites de références. La seconde solution consiste à demander à des spécialistes du domaine de les poser. Ce sont alors eux qui jugent de la pertinence des résultats obtenus (généralement d'après les références et non les documents originaux). Lorsqu'il s'agit de bases plus spécifiques comme la chimie, les évaluateurs peuvent ajouter des questions posées à partir de champs particuliers. On peut citer par exemple le numéro chimique ou le RN (registry number attribué par les chemical abstracts) (41), la recherche par numéro de brevet est également possible dans certaines bases de données (31).

La deuxième étape consiste à préparer les stratégies d'interrogation. Elles le sont à l'aide des manuels et autres documents proposés par le producteur et le serveur. Ces stratégies doivent s'adapter à chaque base de données. Il faut ainsi employer le vocabulaire contrôlé approprié, le but étant d'exploiter le système au maximum de sa capacité.

La dernière partie de l'étude est bien sûr consacrée à la lecture des résultats et à leur interprétation. Répondre à une série de questions que les évaluateurs auront préalablement définies est la méthode couramment employée. On peut proposer ici celles qui sont posées pour une comparaison entre deux bases concernant le marketing. Elles cherchent à évaluer les différences au niveau de la couverture et du nombre de références obtenues (30):

* Les deux bases couvrent-elles la littérature dans la même optique? Trouve-t-on des articles communs dans les deux bases?

* Quelle est l'étendue du chevauchement entre les deux? (Etablissement du pourcentage).

* Quels sont les articles qui n'apparaissent que dans un seul cas? Leur pourcentage? Quelles sont les différences, de ce point de vue, entre les deux bases?

Les résultats peuvent être transcrits dans un tableau. Cette comparaison devient alors une analyse statistique. On y introduit la question, le nom de la base de données, le nombre de références obtenues, le nombre de références uniques dans la base et le nombre de documents communs (lorsqu'on compare seulement deux bases).

Exemple:

A la question de savoir combien il faut interroger de bases de données pour obtenir une recherche la plus complète possible dans le domaine médical K. John répond par un type d'étude proche de celui énoncé précédemment (35).

Il compare huit bases en prêtant une attention toute particulière au choix des termes pour l'interrogation. Il utilise par exemple certains termes génériques pour les noms de médicaments. Mais il estime que l'emploi du vocabulaire contrôlé peut nuire à la comparaison dans la mesure où il diffère dans chaque base. Il cherche à retrouver les articles de certains auteurs dans toutes les bases. La somme de tous les articles trouvés pour chaque auteur est égale à 100 %. Le nombre d'articles trouvés dans une base pour chaque rédacteur est interprété comme une mesure de l'importance de la base et le nombre d'articles non découverts est considéré comme une mesure du risque par omission de la base.

K. John transcrit les résultats obtenus dans un tableau pour les analyser. Il en conclut qu'on ne peut prédire où seront les articles les plus importants. Pour une recherche pointue, il est préférable de mélanger les résultats de plusieurs bases. Le choix doit aussi tenir compte du coût d'interrogation et du temps d'obtention des documents.

Toutes ces comparaisons restent souvent très superficielles, elles donnent quelques points de comparaison qui mettent en évidence les forces et les faiblesses d'un système. Elles se transforment parfois en une sorte de publicité déguisée pour une base précise.

3- Les comparaisons spécifiques: un aspect de la base.

* Comparer la qualité.

Le contrôle de la qualité se conjugue en fait avec la quantité. Dans le domaine médical par exemple l'importance des informations très récentes se révèle souvent essentielle. La qualité se mesure par rapport à la vitesse de mise à jour de la base, aux sources c'est à dire le nombre de journaux dépouillés. L'indexation affecte la recherche et les résultats (48).

La qualité variable entre les bases, se modifie aussi à l'intérieur d'une base à moyen et long terme, une vérification à ce niveau peut se révéler intéressante. L'examen des descripteurs dans le temps s'impose donc comme une mesure régulière.

La qualité ce peut être aussi la mesure combinée de plusieurs facteurs: la couverture des monographies, des périodiques, l'indexation, et une étude résultant des commentaires des utilisateurs. Ce dernier aspect reste peu développé dans les évaluations de ce type. Le spécialiste imagine les difficultés probables du chercheur mais il ne lui demande que rarement sa participation ou uniquement pour le choix des questions (46).

L'évaluation de la qualité ne diffère pas des comparaisons générales quant à la méthodologie employée.

* Comparer l'indexation.

C'est un élément essentiel de la recherche et son évaluation exige la compétence de spécialistes de la question. On notera d'ailleurs que les articles proviennent toujours des mêmes groupes de personnes.

Il s'agit alors d'interroger les bases de données. L'établissement des sujets de recherche est confié à des spécialistes de la question. Chaque sujet est recherché dans plusieurs bases selon plusieurs stratégies: emploi du langage naturel, du langage contrôlé et de références clés. Les résultats sont évalués pour leur pertinence et leur nouveauté. Transcrits dans un tableau, ils sont ensuite discutés.

Plusieurs méthodologies se proposent de vérifier l'indexation d'une base de données.

La procédure la plus courante consiste à effectuer un test de recouvrement. Il prend en compte les jugements pertinents et la satisfaction des utilisateurs mais ces mesures restent subjectives car influencées par les orientations de l'utilisation de la recherche.

Zunde et Dexter proposent, eux, de mesurer la qualité de performance de l'indexeur ou du système indexé (58). La capacité de sélectionner des termes indexés qui donneraient le maximum d'informations sur le document indexé semble facile en théorie. Mais cette réalisation se heurte dans la réalité à deux problèmes majeurs: la détermination d'un groupe universel des termes indexés par document et le degré de concordance parmi les indexeurs pour respecter la signification d'un terme.

La dernière méthode proposée est celle de Griffith et White (55,57). Chaque test consiste à obtenir un groupe d'environ 5 documents connus pour être liés au sujet. Leurs descripteurs sont récupérés. Ils mettent dans un tableau le nombre moyen de descripteurs attachés à tous les documents et à une majorité des documents du groupe. Des statistiques concernant la façon dont chaque base relie les documents en relation et toute une sélection aussi bien grossière que fine des documents sont réalisées. Ils obtiennent aussi une appréciation qualitative dans la description et la pertinence des termes indexés disponibles. Ces évaluations peuvent se référer à des mesures de consistance et d'exhaustivité. Les études de Chu C. et Ajiferuke I. (52,53) prennent en plus en compte la taille de la collection de la base de données tout en conservant la même méthode.

Toutes ces recherches effectuées sur l'indexation conduisent à une seule et même conclusion: plus l'indexation est bonne, plus on a de chance de retrouver le document.

* Comparer la couverture.

La couverture documentaire est loin d'être négligeable lorsqu'on entreprend une recherche. L'orientation de base, son optique, les secteurs couverts peuvent se révéler fondamentaux

Deux méthodes sont actuellement proposées pour évaluer cette couverture documentaire (62):

- * Une méthode bibliographique.
- * Une méthode de profil par sujet.

La première méthode utilise une bibliographie spécialisée ou des articles de synthèse sur un sujet. Cette liste constituée, les éléments sont recherchés dans les bases. Une base qui obtient un fort pourcentage de ces documents est considérée comme ayant une couverture exhaustive. Mais les articles choisis et récupérés n'impliquent aucun jugement de qualité (63).

Cette méthode est considérée comme la méthode standard pour l'estimation de la couverture d'une base en relation avec un domaine spécifique. L'inconvénient majeur de ce système réside dans son coût. Pour obtenir des résultats fiables, le nombre d'articles recherchés doit être suffisant, le nombre d'heures passées en ligne est alors important ce qui augmente le coût. Cette méthode donne donc des informations détaillées sur la façon dont la base couvre l'information sélectionnée à travers les articles.

La méthode du profil doit couvrir une grande partie du sujet (59). Il s'agit d'établir ici une série de questions, de formuler des équations de recherche adaptées à chaque base. La pertinence des documents obtenus est vérifiée par des chercheurs, et les résultats sont entrés dans un tableau. Cette méthode permet d'évaluer la couverture mais aussi le chevauchement des bases. Celle qui obtient le plus de références pertinentes est considérée comme la plus performante sur le sujet. Cette méthode donne une idée plus précise de la spécialisation de la base. Ce système se révèle plus économique en argent et en temps, il est également simple à utiliser. L'emploi de termes différents dans les bases en fonction de leur indexation en est le principal inconvénient.

Les deux méthodes proposées sont totalement fiables et les résultats obtenus par les deux systèmes sont sensiblement équivalents. C'est donc à l'évaluateur d'opter pour l'une ou l'autre méthode en fonction du résultat souhaité.

CONCLUSION.

L'évaluation des bases de données documentaires est devenue un véritable enjeu économique. La parution d'un article très critique peut se traduire dans la réalité par une baisse du nombre d'interrogations. Certains établissements comme la National Library of Medicine aux Etats-Unis ont entrepris une véritable politique d'évaluations. De nombreuses études sont réalisées sur Medline, leur base de données qui est d'ailleurs l'une des plus interrogées dans le monde.

L'évaluation permet de situer une base par rapport à son environnement, mais aussi de déceler les erreurs et de les réparer.

Si les principes généraux sont toujours identiques, les spécialistes affinent leurs études et leurs méthodes au fil du temps. Ces spécialistes deviennent d'ailleurs de plus en plus spécialisés, on peut citer Belver Griffith et Howard White en ce qui concerne l'indexation, Bonnie Snow pour le domaine médical, Frederic Lancaster et Peter Vigil pour les théories générales. Les trois critères clé pour évaluer une base de données: la qualité, l'indexation et la couverture sont l'objet de mesures particulières. Ce sont ces méthodes qui ont le plus évolué au cours de ces dernières années. Les supports ne sont pas non plus négligés dans ces études, les parts de marché du CD Rom ne cessent de croître et la convivialité qu'il propose devient un critère de choix.

Mais il est vrai que tous ces éléments ont un impact direct sur la recherche documentaire entreprise par l'utilisateur. Ces évaluations contribuent donc à faire avancer les recherches automatisées

Des évaluations de stratégie de recherche sont également proposées à l'utilisateur. Elles ne sont pas traitées ici car elles n'entrent pas dans le sujet. Ce type de comparaison tend aussi à se développer. L'effet CD Rom n'y est certainement pas étranger.

III- BIBLIOGRAPHIE.

A- Les supports des bases de données

1- Les serveurs:

1- BUNTROCK, Robert E. and PALMA, Mary Ann. Searching the Beilstein database online : a comparison of systems . *Database* december 1990 , vol. 13 , no 6 , p . 19-34 .

2- FEIGLOS, Susan J . Medline at BRS , Dialog , and NLM : is there a choice . *Bulletin of the medical library association* , 1983, vol. 71 , no 1 , p . 6-12 .

3- HUNTER, Janne A . Medline on BRS , Datastar , and Dialog : does it make a difference ? *Database* , 1991 , vol. 14 , no 1 , p . 43-50 .

4- SNOW, Bonnie. Differences in cancerlit on medlars and dialog. *Online*, november 1986, vol.10, no 6, p. 118-123

5- SNOW, Bonnie. Sedbase online for drug side effect and interaction. *Database*, february 1989, vol. 12, no 1 , p . 85-94.

6- VAN BRAKEL, P.A. Criteria for evaluation of database access software . In the 11th international online information meeting proceedings , London 8-10 dec 1987 ; organised by learned information (Eurpoe) . *Online information 87 : proceedings*. Oxford; Medford : learned informtion , 1987 , p . 111-122 .

2- Les nouveaux supports: les CD Rom:

7- BAKKER, Suzanne; BLEEKER, Ans; VAN DER BURG, Jan; et al . Medline on CD-Rom : a comparison . *Online Review* , February 1989 , vol. 13 , no 1 , p . 39-50 .

8- HARTLEY, R.J. Lisa on CD Rom : an evaluation . *Online review* february 1989 , vol. 13 , no 1 , p . 53-56 .

9- HERTHER, Nancy. A planning model for optical product evaluation. *Online*, september 1986, vol. 10, no 5, p. 128-130.

10- KAYA, Kathryn K and YANG, Andrea M . Consumer health information from information access compagny . *Database* , 1991, vol. 14 , no 1 , p . 54-56 .

11- KLUEGEL, Kathleen and LOEHR, Eric. " Can I find it on païs ? " Comparing Païs on CD Rom and Païs on dialog . *Database* , december 1990 , vol. 13 , no 6 , p . 37-41 .

12- LARGE, J.A. Evaluating online and CD Rom reference sources . *Journal of librarianship*, april 1989, vol. 21 , no 2 , p . 87-108 .

13- MORGAN, V.E. Medical literature searches : online versus CD Rom . *Journal of the Hong kong library association* , 1988 , no 12 , p . 53-57 .

14- POPE, Carslyn. An evaluation of ERIC on CD Rom in a college library. *Electronic Library*, april 1989, vol. 7, no 2, p. 94-97.

15- RAPP, Barbara A; SIEGEL, Elliot R; WOODSMALL, Rose Marie ; LYON-HARTMANN, Becky. Evaluating Medline on CD Rom : an overview of field tests in library and clinical settings. *Online review* , june 1990 , vol. 14 , no 3 , p . 172-186 .

16- REESE, Jean; STEFFEY, Ramona. Eric on CD Rom: a comparison of dialog on disc, OCLC's search CD 450 and Silver platter. *Online*, september 1987, vol. 11, no 5, p. 42-54.

17- ROWLEY, J.E. CD Rom versus online : an evaluation of the effect of the interface on search effectiveness - a pilot study . In the 13th information online information meeting , London 12 dec 1989; organised by learned information (Europe). *Online information 89 : proceedings* . Oxford ; Medford : learned information , 1989 , p . 183-193 .

18- SIEVERTO, E.G.; MASTENBROEK,G.; NIEUWENHUYSEN, P . Information retrieval on a micro-computer : the evaluation of software for personal information systems and for CD-Rom products . In The 12th international online information meeting , London 6-8 dec 1988 ; organised by learned information (Europe) . *Online information 88 : proceedings* . Oxford , Medford : learned information , 1988 , vol 2 , p . 581-595 .

19- SNOW, Bonnie. Life science sources on laserdisk. *Online*, march 1987, vol.11, no 2, p.113-116.

20- TERBILLE, Charles. LISA on CD Rom. *CD Rom librarian*, september 1990, vol. 5, no 8, p. 26-28.

B- L'évaluation du contenu des bases de données.

1- Les théories générales d'évaluation:

21- HARTER, Stephen P . *Online information retrieval : concepts , principles and techniques* . Orlando ; New-York ; London : Academic Press , 1986 . 259 p . Library and information science . ISBN 0-12-328455-4 .

- 22- LANCASTER, F W . *Information retrieval systems : characteristics , testing and evaluation* . 2nd ed . New-York ; Chichester ; Brisbane : J . Wiley and sons , cop 1979 . 381 p . Information sciences series . ISBN 0-471-04673-6 .
- 23- LANCASTER, F.W. *Principes directeurs pour l'évaluation des systèmes d'information* . Unesco . Unisit . Paris : Unesco , 1978 . p.
- 24- MC CARN, Davis B. and LEWIS, Craig M. A mathematical model of retrieval system performance . *Journal of the American society for information science* , 1990 , vol. 41 , no 7 , p . 495-500.
- 25- PAO, Miranda Lee . *Concepts of information retrieval* . Englewood : Libraries unlimited , 1989 . 285 p . ISBN 0-87287-405-2
- 26- RO, Jung soon. An evaluation of the applicability of ranking algorithms to improve the effectiveness of full-text retrieval. I. On the effectiveness of full-text retrieval. *Journal of the American society for information science*, march 1988, vol. 39, no 2, p. 73-78.
- 27- VIGIL, Peter J. Analytical methods for online searching. *Online review*, december 1983, vol. 7, no 6, p. 497-514.
- 28- VIGIL, Peter J. *Online retrieval : analysis and strategy* . New-york ; Chichester ; Brisbane : J . Wiley , 1988 . 242 p . Information sciences series . ISBN 0-471-82423-2 .
- 2- Les comparaisons générales:
- 29- BORKENT, JH; OUKES, F; NOORDIK, JH. Chemical reaction searching compared in REACCS, SYNLIB, and ORAC. *Journal of chemical information and computer sciences*, 1988, vol. 28, no 3, p. 148-150.
- 30- BURSUK, Barbara. A comparison of online marketing and advertising sources : Ami and Mars . *Database* , april 1989 , vol. 12 , no 2 , p . 72-74 .
- 31- DAVIS SMITH, E.M . A comparison of number of patent databases for chemical searches . *International online information meeting* . 9/1985-12-03/ London . p . 437-447 .
- 32- DREYFUS, J; COMBEROUSSE, M. De l'urbanisme à l'urbanite - banques de données et processus d'information. *Bulletin des bibliothèques de France*, 1986, vol.31, no 6, p. 556-563.
- 33- ERNEST, Douglas J. ; LANGE, Holley R. ; HERRING, Della. An online comparison of 3 library science databases. *Reference Quarterly*, winter 1988, vol. 28, no 2, p.185-194.

- 34- FRANSAZOVA, M . Comparative analysis of two metallurgical databases : Viniti-metallurgy and Compadex section metallurgy. In The 12th international online information meeting , London 6-8 dec 1988 ; organised by learned information (Europe) . *Online information 88 : proceedings* . Oxford ; Medford : learned information , 1988 , vol 2 , p . 561-568 .
- 35- JOHN, K . Medical literature searches : how many bibliographic databases are needed for sufficient retrieval in medical topics ? *Methods of information in medicine* , 1985 , vol. 24 , no 3 , p . 84-89 .
- 36- LANSELLE, Marie-Claude . Etude critique des bases et banques de données interrogées par le service de documentation d'une entreprise de batiment et de travaux publics . Paris :INTD , 1983 .
- 37- MEURLING, Anita . Cas online and Darc : a comparison . *Database* , february 1990 , vol. 13 ,no 1 , p . 54-63 .
- 38- MONTGOMERY, Susan. British databases : a comparison of Reuter textline and Profile information . *Database* , april 1990 , vol. 13 , no 2 , p . 28-37 .
- 39- NIXON, Judith M. Online searching for human nutrition: an evaluation of databases. *Medical reference services quarterly*, 1989, vol. 8, no 3, p. 27-35.
- 40- PERDUE, Bob. Biosis Previews: implication for research in behavioral medicine. *Medical reference services quarterly*, 1988, vol.7, no 2, p. 57-62.
- 41- SANOUILLET, A . Francis : a comparison with Eric and Sociological abstracts . In the 11th international online information meeting proceedings , London 8-10 dec 1987 ; organised by learned information (Europe) . *Online information 87 : proceedings* . Oxford ; Medford : learned information , 1987 , p .
- 42- SNOW, Bonnie . Health periodicals databases : a closer look . *Database* , 1991 , vol.14 ,no 1 , p . 84-89 .
- 43- STIEG, MF; ATKINSON, JL. Librarianship online: old problems, no new solutions. *Library journal*, 1988, vol. 113, no 16, p. 48-59.
- 44- WOOD, Frances E; BERRIE, Andrew T; PAMPLIN, Helen R; et al. Evaluations using test queries of chemical hazards databases and databanks . *Journal of information science* , 1989 , vol. 15 , no 4-5 , p . 269-276 .

3- Les comparaisons spécifiques: un aspect de la base

45- MC CAIN, K.W. ; WHITE, Howard D. ; GRIFFITH, B.C. Test retrieval as a measure of system performance: Medline and the medical behavioral sciences. In : *Asis'86. Proceedings of the 49th Asis annual meeting*, vol. 23, Chicago, Illinois, september 28- october 2, 1986. Edited by Julie M. Hurd and Charles H. Davies, Medford, New Jersey, learned information inc. for the american society for information science, p. 199-203.

46- GRIFFITH, Belver C; WHITE, Howard D; CARL DROTT, M; SAYE, Jerry D. Tests of methods for evaluating bibliographic databases: an analysis of the National Library of Medicine's handling of literatures in the medical behavioral sciences. *Journal of the American society for information science*, july 1986, vol. 37, no 4, p. 261-270.

* La qualité:

47- GORDON, MD. Evaluating the effectiveness of information retrieval systems using simulated queries. *Journal of the American society for information science*, july 1990, vol. 41, no 5, p.313-323.

48- SOREMARK, Gun . Medline versus Embase : comparing search quality . *Database* , December 1990 , vol. 13 , no 6 , p . 66-67 .

49- TODESCHINI, C; FARRELL, MP. An expert system for quality control in bibliographic databases. *Journal of the American society for information science*, january 1989, vol. 40, no 1, p. 1-11.

* L'indexation:

50- BOYCE, Bert R. and MC LAIN, John P. Entry point depth and search using a controlled vocabulary . *Journal of the American society for information science* , 1989 , vol. 40 , no 4 p . 273-276 .

51- BRENNER, Sandra H. ; MC KININ, Emma Jean. CINAHL and Medline: a comparison of indexing practices. *Bulletin of the medical library association*, october 1989, vol. 77, no 4, p. 366-371.

52- CHU, CM; AJIFERUKE, I. Quality of indexing in online databases: an alternative measure for a term discriminating index. *Information processing and management*, 1988, vol. 24, no 5, p.599-601.

53- CHU, Clara M. and AJIFERUKE, Isola . Quality of indexing in library and information science databases . *Online Review* , February 1989 , vol. 13 , no 1 , p .11-22 .

54- MAC CAIN, Katherine. Descriptor and citation in the medical behavioral sciences literature : retrieval overlaps and novelty distribution . *Journal of the American society for information science* , 1989 , vol . 40 , no 2 , p .110-114 .

55- MAC CAIN, K.W; WHITE, H.D; GRIFFITH, B.C. Comparing retrieval performance in online databases. *Information processing and management*, 1987, vol. 23, no 6, p. 539-553.

56- SIEVERT, M. and VERBECK, A. The indexing of the literature of online searching: a comparison of Eric and Lisa. *Online review*, april 1987, vol. 11, no 2, p. 95-104.

57- WHITE, Howard D. and GRIFFITH, Belver C. Quality of indexing in online databases . *Information Processing and management* , 1987 , vol. 23 , no 3 , p . 211-224 .

58- ZUNDE, P. and DEXTER, M.E. Indexing consistency and quality. *American documentation*, 1969, 20, p. 259-267.

* La couverture:

59- BRANDSMA, R; DEURENBERG-VOS, H.W.J; BAKKER, S; et al . A comparison of the coverage of clinical medicine provided by Biosis Previews and Medline . *Online review* , december 1990 , vol. 14 , no 6 , p . 367-375 .

60- BROOKS, K. A comparison of the coverage of agricultural and forestry literature on Agricola, Biosis, Cab and Scisearch. *Database*, march 1980, vol. 3, no 2, p. 38-49.

61- KELLY, Sarah A. Retrieving information on the neurophysiology of speech . *Medical references services quarterly* , 1988 , vol. 7 , p . 31-45 .

62- TENOPIR, C. Evaluation of database coverage: a comparison of two methodologies. *Online Review*, 1982, vol. 6, no 5, p. 423-441.

63- YONKER, Valerie A. ; PENNY YOUNG, K. ; HORWITZ, Seth ; COUSIN, Karen. Coverage and overlaps in bibliographic databases revelant to forensic medicine: a comparative analysis of Medline. *Bulletin of the medical library association*, january 1990, vol. 78, no 1, p. 49-56.

C- Les répertoires

64- ADBS; ANRT. *Répertoire des banques de données professionnelles 1990*. 12ème ed. Paris: ADBS, 1990. ISBN 2-901046-29-00.

65- FARBEY, Roger. *Medical databases 1988*. 4th ed. London: Aslib, 1987. 88 p. An Aslib online guide. ISBN 0-85142-226-8.

66- HALL, James H . *Online bibliographic databases : a directory and sourcebook* . 4th edition . London : Aslib , 1986 . 509 p . ISBN 0-85142-202-0 . ISBN 2-8103-2080-0 .

1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025

1945

1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025

1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025

1945

1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025

1945

1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025

1945
1946
1947
1948
1949
1950
1951
1952
1953
1954
1955
1956
1957
1958
1959
1960
1961
1962
1963
1964
1965
1966
1967
1968
1969
1970
1971
1972
1973
1974
1975
1976
1977
1978
1979
1980
1981
1982
1983
1984
1985
1986
1987
1988
1989
1990
1991
1992
1993
1994
1995
1996
1997
1998
1999
2000
2001
2002
2003
2004
2005
2006
2007
2008
2009
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
2020
2021
2022
2023
2024
2025

1945