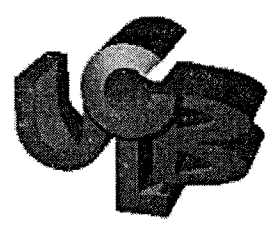




ensib
Ecole Nationale Supérieure
des Sciences de l'Information
et des Bibliothèques



Université
Claude Bernard
Lyon I

DESS Informatique Documentaire

Rapport de recherche bibliographique

**Le catalogage informatisé des documents en écritures
non alphabétiques : cas du chinois, du japonais et du coréen**

Chisato SUGITA

Sous la direction de

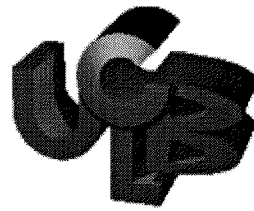
Monsieur Jean-Louis BOULLY

Fonds chinois - Bibliothèque municipale de Lyon

Année 1997-1998



enssib
Ecole Nationale Supérieure
des Sciences de l'Information
et des Bibliothèques



Université
Claude Bernard
Lyon I

DESS Informatique Documentaire
Rapport de recherche bibliographique

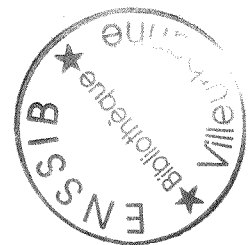
**Le catalogage informatisé des documents en écritures
non alphabétiques : cas du chinois, du japonais et du coréen**

Chisato SUGITA

Sous la direction de

Monsieur Jean-Louis BOULLY

Fonds chinois - Bibliothèque municipale de Lyon



Année 1997-1998

Le catalogage informatisé des documents en écritures non alphabétiques : cas du chinois, du japonais et du coréen

Chisato SUGITA

RESUME

Ce rapport a pour objet la description des caractéristiques de l'environnement informatique dans le cas des écritures chinoise, japonaise et coréenne (CJK). Nous étudierons ici les problèmes liés au codage des caractères et évoquerons la norme Unicode qui pourra sans doute apporter des solutions satisfaisantes. Toutefois, actuellement des solutions existent déjà qui permettent l'intégration des caractères CJK.

DESCRIPTEUR

Catalogage ; CJK ; Chinois ; Japonais ; Coréen ; Unicode

ABSTRACT

This report aims at describing the special features of the Chinese, Japanese, and Korean (CJK) record environment. We study here the issues concerned with the coding of character records with reference to the Unicode standard that could solve the problem in a satisfactory way. At the present moment is already possible to solve the problem of integrating the CJK characters.

KEYWORD

Cataloging ; CJK ; Chinese ; Japanese ; Korean ; Unicode

INTRODUCTION	1
MÉTHODOLOGIE	2
1 PREMIÈRE APPROCHE DU SUJET ET SÉLECTION DE MOTS-CLÉS	2
2 RECHERCHE SUR BASES DE DONNÉES.....	2
2-1 Consultation du catalogue de l'ENSSIB.....	2
2-2 Consultation des bases de données sous forme de CD-ROM	3
2-3 Interrogation des bases de données en ligne via DIALOG.....	5
2-4 Interrogation des bases de données en ligne via STN International	8
2-5 Recherche sur Internet.....	9
Conclusion	14
SYNTHÈSE	16
AVANT PROPOS	16
1 CARACTÈRES CJK.....	16
2 CARACTÉRISTIQUES DE L'UNICODE ET DE LA NORME ISO/IEC10646.....	17
3 LES SYSTÈMES DES GRANDS RÉSEAUX DE BIBLIOTHÈQUES POCÉDANT LES EXTRÊME-ORIENTAUX	18
3-1 OCLC (<i>Online Computer Library Center</i>).....	18
3-2 RLIN (<i>Research Libraries Information Network</i>).....	19
3-3 UTLAS (<i>University of Toronto Library Automated System</i>).....	19
3-4 NLA (<i>National Library of Australia</i>)	19
3-5 CNRS (<i>Centre National de la Recherche Scientifique en France</i>)	20
4 PROBLÈMES	20
5 CONCLUSION.....	21
ANNEXE 1	22
LA LISTE DES FIRMES ET DES BIBLIOTHÈQUES PROPOSANT DES SYSTÈMES DE GESTION AUTOMATISÉE DES BASES DE DONNÉES MULTI-LINGUES AVEC POSSIBILITÉS DE MULTISCRIPTE CJK.	
ANNEXE 2	24
LISTE DES LOGICIELS QUI GÉNÈRENT LES CARACTÈRES CJK.	
ANNEXE 3	25
EXEMPLE DU LOGICIEL WINMASS	
ANNEXE 4	26
EXEMPLE DU SYSTÈME CJK PLUS PAR L'OCLC : RÉSULTAT DE RECHERCHE PAR NOM D'AUTEUR	

JAPONAIS.

BIBLIOGRAPHIE	27
1 - OUVRAGES	27
2 - ARTICLES DE PERIODIQUES ET ARTICLES ELECTRONIQUES TROUVES SUR LE WEB	27
3 - RAPPORTS DE CONFERENCE.....	31
4 - BREVETS.....	31
5 - SITES INTERNET	31

Introduction

L'automatisation informatique des bibliothèques se développe de plus en plus. Un réseau informatisé de bibliothèques consiste à faire travailler en ligne des bibliothèques autour d'une base de données bibliographiques commune. Cela ainsi que nous avons trouvé de très nombreux systèmes qui offrent des versions incluant les graphiques vernaculaires et leurs transcriptions.

Dans le cas des caractères latins, l'accentuation constitue la différence la plus importante : français, allemand, espagnol, etc..

Pour les régions du monde qui n'utilisent pas l'alphabet latin, des difficultés existent également pour faire apparaître correctement certains caractères des écritures alphabétiques par exemple.

Grâce à la communication informatique développée comme Internet, la demande d'échange des catalogues ou la gestion des bases de données bibliographiques en multilingue est en constante augmentation, tout en respectant la langue originale.

Aux Etats-Unis, les bibliographic utilities, (serveurs bibliographiques) offrent à un ensemble de bibliothèques un système de catalogage partagé, ainsi qu'un système de prêt-inter-bibliothèques.

C'est pourquoi il faut normaliser la gestion des systèmes informatiques et prendre rencontre les différences des chacune des écritures.

Nous pouvons accéder aux données en caractères multi-script, mais la lecture des caractères est actuellement soumise aux environnements informatiques nationaux. Notamment, la lecture des caractères de chinois, japonais et coréen (CJK), sur l'écran français. Dans ce cas-là, nous transférons des fichiers en romanisé et ne recevons pas toujours les notices CJK directement. Grâce au développement du réseau Internet, nous pouvons facilement accéder à des sites monde entier. Mais il n'est pas toujours facile de prendre connaissance des informations en écriture vernaculaire. Nous nous sommes donc notamment intéressés au codage, qui permet de visualiser ou de saisir des caractères CJK directement. Nous avons donc choisi ce sujet sur le système CJK et recherché ses caractéristiques et ses fonctions dans plusieurs réseaux existants.

Nous allons également traiter de la norme Unicode, concernant le CJK, car elle semble en théorie être celle qui résoudrait tous ces problèmes.

Méthodologie

1 Première approche du sujet et sélection de mots-clés

Au début de notre recherche, nous avons commencé par interroger l'OPAC de la bibliothèque de l'ENSSIB, afin d'établir les mots-clés pour la recherche. Aussi plusieurs consultations avec notre commanditaire nous ont permis de préciser le sujet et de choisir la démarche.

Afin d'éviter les nombreux bruits, nous avons sélectionné les descripteurs par l'étape de première approche : «CJK», «CJK350», «CJKPLUS» développé par CJK, «UNICODE», «ISO10646» (la norme pour multilingue), «RLIN», «UTLAS», «OCLC» (les réseaux diffusent le système CJK), etc.. Nous avons aussi décidé de limiter l'année de publication après 1993.

Chaque recherche est préalablement soumise au dictionnaire ou l'index des bases de données consultées afin d'éviter les mots n'ont pas été retenus.

2 Recherche sur bases de données

2-1 Consultation du catalogue de l'ENSSIB

Recherche multi-champs (par mots du titre)

Équations de recherches	Résultat	Résultat pertinent
gestion catalogue bibliothèque	3	0
catalogue informatisé	1	0
logiciel documentaire	24	0
automated catalog	3	0
automated system	7	2
informatique bibliothèque	20	1
OCLC catalogue	2	0
RLIN	0	-
CJK	0	-
multilingue	3	0
multiscript	2	1
ISO 10646	0	-
unicode	1	1

2-2 Consultation des bases de données sous forme de CD-ROM

◆ Consultation du CD-ROM **LISA-PLUS**

LISA (Library and Information Science Abstracts Plus)-base spécialisée en sciences d'information et bibliothéconomie, établie par the Library Association et par ASLIB (deux associations professionnelles anglaises) à partir de tous types de documents. L'interrogation de DIALOG en ligne mise à jour est aussi nécessaire, elle a été réalisée plus loin (cf. 2-3).

Recherche par Expert Search

ab = abstract

cs = Combine Search

da = Publication Date

kt = title keyword

kw = keyword

su = subject

ti = title

Équations de recherche	Résultat	Résultat pertinent
kw="RLIN"	259	
kw="asiagraphics"	1	0
kw="RLIN" AND (da="1993" OR da="1994" OR da="1995" OR da="1996" OR da="1997")	49	
kw="RLIN" AND (da="1993" OR da="1994" OR da="1995" OR da="1996" OR da="1997") AND ("CJK" OR "CJK 350")	3	3
kt="RLIN" AND (da="1993" OR da="1994" OR da="1995" OR da="1996" OR da="1997")	17	2
kt="CJK"	20	
kw="CJK"	39	
kt="CJK350"	1	1
kw="CJK350"	4	4
su="CJK" OR "CJK Plus"	2	2
ab="CJK" OR "CJK 350"	11	11
ti="CJK"	3	3
(kt="CJK" OR kt="CJK350") AND (da="1993" OR	4	4

Équations de recherche	Résultat	Résultat pertinent
da="1994" OR da="1995" OR da="1996" OR da="1997")		
(kw="CJK" OR kw="CJK350") AND (da="1993" OR da="1994" OR da="1995" OR da="1996" OR da="1997")	11	7
kt="unicode"	3	1
kw="unicode"	14	2

◆ Consultation du CD-ROM **PASCAL**

Pascal - base multilingue (français, anglais, espagnol) qui couvre la littérature internationale (périodiques, ouvrages, comptes-rendus, thèses, rapports) en Sciences, Technologie et Médecine, établie par l'INIST (Institut national de l'information scientifique et technique). Il est aussi disponible en ligne par DIALOG et nous avons consulté (cf. 2-3).

DXA = pour un mot composé ou une expression en anglais

DXF = pour un mot composé ou une expression en français

* = troncature pour trouver une expression, il permet d'utiliser le dictionnaire

Équations de recherche	Résultat	Résultat pertinent
DXA="ISO 10646" ET "NONLATIN"	2	2
DXA="catalog*" ET "libra*" ET "automation"	14	0
DXA="automated catalog"	0	-
DXA="catalog*" ET "ibra*" ET "automation" ET "multiling*"	2	2

◆ Consultation du CD-ROM **INSPEC**

INSPEC- base qui couvre les domaines de la physique, de l'électronique, de l'ingénierie nucléaire, de l'ingénierie électrique, de l'informatique, de l'intelligence artificielle, de la technologie de l'information, ect.. Il est disponible à Doc'INSA et permet d'une recherche par adjacence pure, les termes seront recherchés côte à côte et dans le sens indiqué. Il est mise à jour trimestriellement et nous avons consulté de Jan. 1994 à Sep 1997. Il est également disponible en ligne par DIALOG. Le résultat de notre recherche par DIALOG a été réalisés après (cf. 2-3).

? = troncature

Équations de recherche	Résultat	Résultat pertinent
catalog? AND (Chinese OR Japanese OR Korean)	22	6
(unicode OR cod?) AND (CJK OR (Chinese OR Japanese OR Korean))	303	
(multiling? OR multiscrypt?) AND (Chinese OR Japanese OR Korean)	27	3
RLIN	17	1
RLIN AND CJK	1	1

Nous avons interrogé les autres CD-ROM: **DOCTHESE; FRANCIS; BNF**, mais les résultats ne sont pas intéressants.

2-3 Interrogation des bases de données en ligne via DIALOG

Le serveur DIALOG donne accès à 450 banques de données internationales et nord-américaines dans les domaines Scientifiques, Techniques, Economiques, Brevets et Marques.

Afin de trouver les bases de données adaptées à notre recherche, nous avons utilisé ONE SEARCH, proposé par DIALOG, qui est une base de données prédéfinie la catégorie par sujets. Nous avons choisi la catégorie «Library and Information Science» qui s'appelle «INFOSCI». Il nous a donné dix bases de données suivantes:

File no. 1	: ERIC
File no. 2	: INSPEC
File no. 6	: NTIS
File no. 61	: LISA
File no. 121	: British Education Index
File no. 148	: IAC Trade & Industry Database
File no. 202	: Information Science Abstracts
File no. 211	: IAC Newsearch
File no.437	: Education Abstracts
File no. 438	: Library Literature

Nous avons choisi les bases de données ci-dessous:

IAC Newsearch- index des articles de journaux et périodiques internationaux.

Information Science Abstracts- sciences de l'information.

INSPEC- Physique, électricité, électronique, informatique: information technique et scientifique. Couverture mondiale. (support CD-ROM disponibilité à DocINSA)

LISA (Library & Information Science Abstracts)- sciences de l'information et bibliothéconomie. (support CD-ROM disponibilité au réseau de l'ENSSIB)

NTIS- Études scientifiques subventionnés par le gouvernement des États-Unis: résultats et applications.

PASCAL- Tous les domaines des sciences et techniques: biologie, médecine, chimie, physique, géophysique, biomédical. (support CD-ROM disponibilité au réseau de l'ENSSIB)

W = opérateur proximité pour préciser l'ordre entre deux mots

? = troncature qui remplace plusieurs caractères

? ? = troncature qui remplace 0 ou 1 caractère

Les questions et les résultats de recherche sont les suivants (comme des résultats nuls n'indiquent pas):

?ss (CJK OR (CJK(w)350) OR (CJK(w)PLUS)) AND py>1993

Base	Nombre de références obtenues	Nombre de références pertinentes	Bruits (%)
IAC Newsearch	3	1	67
INSPEC	4	3	25
ISA	3	3	0
LISA	7	5	29
NTIS	4	1	75
PASCAL	3	3	0

?ss (UNICODE OR (ISO(w)10646)) AND py>1993

Base	Nombre de références obtenues	Nombre de références pertinentes	Bruits (%)
IAC Newsearch	26	1	96
INSPEC	27	2	92
ISA	3	2	33
LISA	12	2	83
PASCAL	6	2	67

?ss catalog? AND automation? ? AND (CJK OR (CJK(w)350) OR (CJK(w)PLUS)
AND py>1993

Base	Nombre de références obtenues	Nombre de références pertinentes	Bruits (%)
IAC Newsearch	2	1	50
INSPEC	1	1	0
ISA	1	1	0
LISA	1	1	0
PASCAL	7	3	57

Après la consultation de certaines bases de données sélectionnées par ONE SEARCH, nous avons interrogé la fiche technique (blue sheet), car le catalogue de serveur n'est pas récent (printemps 1996). La fiche permet la liste des dernières bases de données disponibles sur le serveur. Nous avons donc sélectionné deux bases japonaises sur l'information scientifique et technique élaborés par le Japan Information Center of Science and Technology:

File no.94 : JICST-Eplus (Japan Information Center of Science and Technology)

File no.347 : JAPIO (Japan Patent Information Organization)

Les questions et les résultats de recherche sont les suivants:

?ss ((Chinese AND Japanese AND Korean) OR CJK) AND py>1993

Base	Nombre de références obtenues	Nombre de références pertinentes	Bruits (%)
JICST-EPlus	39	8	79
JAPIO	4	3	25

?ss ((Chinese AND Japanese AND Korean) OR CJK) AND py>1993 AND (software
OR database)

Base	Nombre de références obtenues	Nombre de références pertinentes	Bruits (%)
JICST-EPlus	1	1	0

?ss ((Chinese AND Japanese AND Korean) OR CJK) AND py>1993 AND catalog?

Base	Nombre de références obtenues	Nombre de références pertinentes	Bruits (%)
JICST-EPlus	2	2	0

Le coût et le temps d'interrogation par source consultée:

\$0.02 0.009 Hrs Estimated cost OneSearch

\$0.92 0.061 Hrs Estimated cost File 2

\$2.25 0.150 Hrs Estimated cost File 6

\$2.13 0.142 Hrs Estimated cost File 61

\$3.00 0.200 Hrs Estimated cost File 94

\$1.58 0.105 Hrs Estimated cost File 144

\$2.10 0.140 Hrs Estimated cost File 202

\$3.24 0.216 Hrs Estimated cost File 211

\$1.50 0.100 Hrs Estimated cost File 347

\$4.3 INTERNET

\$21.20 Estimated cost this search

\$44.83 Estimated total session cost 2.342 Hrs

Le coût total d'interrogation de DIALOG dont la durée est 2 h 21 mn. est égal à \$44.83 (environ 273 Frs).

2-4 Interrogation des bases de données en ligne via STN International

Le serveur STN International (Scientific & Technical Information Network) est un des grands services en ligne d'information scientifique et technique. Il est le fruit d'une collaboration entre le FIZ (Fachinformationszentrum) Karlsruhe, le CAS (Chemical Abstracts Service) et le JICST (The Japan Science and Technology Corporation, Information Center for Science and Technology). Il est possible d'accéder aux 112 bases de données sur tous les aspects des sciences et de la technique. Il permet aussi d'accéder à certains bases de données par Internet en payant (login et password).

Nous avons recherché par les catégories "General Science", "Math/Computer Science", elles nous permettent d'accéder facilement aux bases de données correspondants. Nous avons interrogé :

CONF , produit par FIZ Karlsruhe (Allemagne), qui contient les informations sur science et technologie.

SCISEARCH , produit par Institute for Scientific Information (Etats-Unis), qui

contient les domaines des sciences et des technologies.

COMPUSCIENCE , produit par FIZ Karlsruhe (Allemagne),qui couvre le domaine des informatiques théoriques et graphiques, des logiciels ou systèmes informatiques.

Les questions (les mots passés par un index proposé) et les résultats sont les suivants :

Question	CJK			CJK AND catalog OR catalogage
	CONF	SCISEARCH	COMPUSCIENCE	COMPUSCIENCE
Base				
Nombre de références	1	0	3	1
Références pertinentes	1	-	3	1
Taux de pertinence (%)	100	-	100	100

Le coût par recherche est 11,70Frs. , l’affichage par notice est 15,70 Frs.

Le coût total de cette interrogation en ligne est 137 Francs (5 recherches et 5 notices affichés) et le temps de consultation est environ 30 mn.

2-5 Recherche sur Internet

2-5-1 Site de l’ENSSIB

Bibliothèque de l’ENSSIB

<http://enssibd2.enssib.fr/doris_web/DorisWelcome.html>

La consultation de la bibliothèque de l’ENSSIB nous a permis d’accéder aux sites nationaux et internationaux qui nous intéressent:

- La British Library <<http://www.portico.bl.uk>>

The British Library a un fonds oriental et indien qui offre un catalogue en CJK.

- La Library of Congress <<http://www.loc.gov/>>

Ses fonds sont riches et y compris des fonds CJK.

- National Library of Australia <<http://www.nal.gov.au/>>

Elle a mission de fournir les notices de ses fonds spéciaux (chinois, japonais et coréen) aux bibliothèques australiennes.

- CEAL (Council on East Asian Libraries)

<<http://darkwing.uoregon.edu/~felsing/ceal/welcome.html>>

Il fournit beaucoup de sources qui correspondent à notre recherche.

2-5-2 Base de données de l'OCLC

OCLC <<http://www.ref.uk.oclc.org:2000>>

OCLC nous a donné la possibilité de tester les bases de données FIRST SEARCH, qui permet l'accès à une soixantaine de bases de données en ligne.

Donc nous avons consulté certaines bases et les résultats sont les suivants:

sujet: CJK, CJK350, CJK PLUS

PapersFirst : liste de rapports de conférence

résultat: 2 résultat pertinent: 2

WorldCat : livres et autres matériels dans le domaine de bibliothèques.

résultat: 20 résultat pertinent: 4

NetFirst : OCLC bases de données sur Internet

résultat: 3 résultat pertinent: 3

LibraryLiterature : matériels de bibliothèques et réseau de bibliothèques

résultat: 2 résultat pertinent: 2

2-5-3 News

Les groupes de News présentent un difficile problème d'exploitation sauf à se contenter d'un petit nombre.

DéjàNews <http://www.dejanews.com/home_ps.shtml>

Il propose une autre approche en considérant les groupes de News comme des bases de données.

Équations de recherche	Résultat	Résultat pertinent
CJK AND catalog	3	3
CJK AND UNICODE	35	3

SIFT <<http://www.reference.com/>>

Il est un service gratuit de filtrage d'informations extraites des groupes de News.

Équations de recherche	Résultat	Résultat pertinent
CJK	48	2
CJK AND UNICODE	2	2

2-5-4 Moteurs de recherche

Yahoo! <<http://www.yahoo.com>>

Yahoo a été récompensé au printemps '95. C'est un catalogue très complet et organisé par thèmes. Il permet aussi la recherche avancée.

Équations de recherche	Résultat	Résultat pertinent
CJK	2	1
"CJK" "software"	1	1
"CJK database"	68	27
"CJK" AND "UNICODE"	3802	
"CJK" "leader"	6756	

Équations de recherche	Résultat	Résultat pertinent
"CJK" "bibliographical database"	4286	
"CJK" "catalog"	13778	

Yahoo!France <<http://www.yahoo.fr>>

En français vers des sites francophones

Équations de recherche	Résultat	Résultat pertinent
"CJK"	3526	
"CJK" AND "logiciel"	53	
"CJK" AND "base de données bibliographique"	3423	

Infoseek

<http://guide.infoseek.com/Home?>

Moteur de recherche performant et rapide. Il indexe les serveurs Web. Il permet de rechercher l'information dans les serveurs Web, Gopher, FTP, les groupes de News et des sites évalués.

Équations de recherche	Résultat	Résultat pertinent
"CJK"	1726	
"catalog" (search only within these 1726 pages)	2	2
"software"	4 427 522	
"CJK" (search only within these 4 427 522 pages)	19	4
"UNICODE"	1 170	

Équations de recherche	Résultat	Résultat pertinent
"CJK" (search only within these 1 170 pages)	60	
"east asian"	17 483	
"CJK" (search only within these 17 483 pages)	623	
"catalog" (search only within these 623 pages)	1	1

Alta Vista

<http://www.altavista.digital.com>

Altavista nous permet d'accéder à l'index le plus important du Web. Il offre deux modes de recherche :

Recherche simple : Il suffit d'entrer les mots les uns à la suite des autres.

Recherche avancée : Nous avons le choix entre les serveurs Web et les groupes USENET. Il faut obligatoirement utiliser les opérateurs AND, OR, NOR ou NEAR pour combiner plusieurs termes.

" " recherche de phrases

* troncatures

title : cheminement de recherche

Nous avons aussi limité les documents mis à jour après le 01 janvier 1993.

Équations de recherche	Résultat	Résultat pertinent
((Japanese AND Chinese AND Korean) OR "CJK") NEAR "bibliographic* AND ("software" OR "database")	424	
title :(Japanese AND Chinese AND Korean) OR "CJK") AND "bibliographic* AND ("software" OR "database")	118	
title :(Japanese AND Chinese AND Korean) OR "CJK") NEAR	3	3

Équations de recherche	Résultat	Résultat pertinent
"bibliographic* AND ("software" OR "database")		
((Japanese AND Chinese AND Korean) OR "CJK") NEAR "bibliographic* NEAR ("software" OR "database")"	34	16
title :("UNICODE" OR "ISO NEAR 10646") AND "CJK"	1	1
"catalog*" NEAR "CJK"	134	
title : "catalog*" NEAR "CJK"	3	3
"CJK bibliographic"	11	7
"CJK bibliographic record"	10	7
title : "CJK" NEAR "RLIN"	3	3
title : "CJK" OR "RLIN CJK"	17	5
"UTLAS" AND "Japanese"	21	3

Conclusion

Le résultat de l'interrogation du catalogue de l'ENSSIB, n'était pas très pertinent. Ainsi nous avons consulté les CD-ROM disponibles sur le réseau de l'ENSSIB et à DocINSA. Ces résultats nous ont permis de sélectionner les concepts ainsi que les mots-clés nécessaires à notre recherche ultérieure.

Pour les interrogations des bases de données sur les serveurs (Dialog et STN), il est indispensable d'identifier, de sélectionner les bases de données, et de choisir les mots-clés lors de la recherche. Nous avons sélectionné certaines bases que nous avons consultées sur CD-ROM, mais les résultats n'étaient pas les mêmes, en raison de mise à jour.

Le temps passé à la recherche d'information sur les serveur en ligne a pu être évalué, mais pour les autres, cela n'a pas été précisé. En ce qui concerne le serveur STN International, nous avons pas obtenu assez d'informations, par exemple le fichier technique n'est pas disponible à la bibliothèque de l'ENSSIB.

Notre stratégie de recherche sur Internet est de compléter les informations non seulement documentaires, mais aussi commerciales.

La recherche via le moteur Yahoo fut difficile en raison du grand nombre de réponses.

Au contraire, les interrogations sur Altavista et Infoseek nous paraissent être les plus efficaces.

La recherche générale sur Internet a pris beaucoup de temps et quelquefois, nous nous sommes perdus et avons abouti à des sites inintéressants.

Afin de rédiger la synthèse, nous avons utilisé ces deux derniers moteurs et compte des résultats pertinents.

Synthèse

Avant propos

Le système du catalogue informatique en multi-lingue des bibliothèques **CJK** (Chinese, Japanese, Korean) fut créé en 1986 par le réseau OCLC (Online Computer Library Center aux Etats-Unis). Les informations bibliographiques sont saisies deux fois, la première en langue originale, la seconde en translittération.[2, 66]¹ Il est donc possible de visualiser les caractères sur l'écran dans l'environnement romanisé.

Actuellement, les différents réseaux ont développé à leurs propres possibilités de générer les écritures chinoise, japonaise et coréenne.

1 Caractères CJK

En ce qui concerne les caractères **CJK**, ils dépendent beaucoup des polices présentes dans les environnements.

En outre, le **CJK** partage un grand nombre de caractères idéographiques d'origine chinoise pour la plupart. Nous les appelons en chinois *Hanzi*, en japonais *Kanji*, et en coréen *hanja*. Il existe plusieurs systèmes de codages des caractères chinois mis au point en Chine populaire et à Taïwan : GB signifie "Guo Biao", qui est un standard chinois et utilisé en Chine et à Singapour, "Big-5" est développé à Taïwan et utilisé à Hong Kong et aussi aux Etats-Unis. Il existe un codage pour le coréen qui s'appelle KSC, est une abréviation par "Korean Standard Character set" le codage plus utilisé en Corée. KSC comporte les caractères chinois et les combinaisons syllabiques possibles de son alphabet *Han'gul*. Le codage des caractères japonais appelle JIS², qui a des capacités de traiter les alphabets latin, grec et cyrillique[18]. (voir la figure 1[64])

Les caractères chinois sont utilisés également dans les écritures japonaise et coréenne, mais des différences sont apparues au cours du temps qui les ont modifiés. Donc l'unification des caractères et la localisation par script posent des difficultés.

¹ N.B. le chiffre mis entre [] renvoie à la référence bibliographique

² JIS est une abréviation par "Japanese Industrial Standard", développé par Microsoft.

Chinois :

Système d'écriture	Caractères idéographiques (traditionnels, simplifiés)
Jeux de caractères	GB2312-80, GBK, GB 13000-1 (sous ensemble de ISO/IEC 10646-1), BIG-5, CNS
Codage	ISO-2022, EUC, HZ ³ , zW, UTF-8, UCS-2

Japonais :

Système d'écriture	<ul style="list-style-type: none"> • un alphabet syllabique écrit en deux styles : <ul style="list-style-type: none"> - style normal (<i>Hiragana</i>) - style spécifique (<i>Katakana</i>) • un certain nombre de caractères idéographiques d'origine chinoise ou proprement japonaise (<i>Kanji</i>)
Jeux de caractères	JIS, JIS 0221-1995 (sous ensemble de ISO/IEC 10646-1)
Codage	ISO-2022-JP, EUC ⁴ , Shift-JIS ⁵

Coréen :

Système d'écriture	Alphabet coréen (<i>Hangul</i>) et, optionnellement, un certain nombre de caractères idéographiques chinois (<i>Hanja</i>)
Jeux de caractères	KSC5600, KSC5000-95 (sous ensemble de ISO/IEC 10646-1) KSC5700-96 (sous ensemble de Unicode 2.0)
Codage	ISO-2022-KR, EUC, <i>Johab</i> (combinatoire), n-Bytes, <i>Wansung</i> (pré-composé)

<Figure 1>

2 Caractéristiques de l'UNICODE et de la norme ISO/IEC10646

Le codage des caractères effectué jusqu'à présent par le standard ASCII est basé sur une représentation binaire utilisant 8 bits et n'autorise que 128 positions de codage. L'ASCII n'exploite que l'alphabet latin.

L'UNICODE fut créé en 1989 et il est un sous ensemble de la norme où les caractères sont codés sur 2 octets. (16 bits, autrement dit the Base Multilingual Plane ou BMP⁶) Le but d'Unicode est de fournir un codage non-ambigu, fixe sur 16 bits, qui n'a besoin ni

³ HZ est une abréviation par "Hanzi", veut dire des caractères chinois, utilisé dans E-Mail ou Newsgroups.

⁴ EUC signifie "Extended UNIX Code", il est populaire dans PCs.

⁵ On trouve Shift-JIS au support MS-DOS, Windows et Macintosh.

⁶ Le BMP est le seul plan qui soit actuellement défini pour ISO/IEC 10646.

de séquences de contrôle, ni de méthodes de compactage. Il doit permettre l'échange, le traitement et la visualisation des caractères du monde entier. Donc Unicode est comme une généralisation à double largeur d'ASCII. Mais Unicode n'est pas parfait pour visualiser les caractères sans logiciel approprié. [38, 40]

La norme ISO/IEC10646 parue en 1993, définit un ensemble universel de codage de caractères sur plusieurs octets (Universal Multiple-Octet Coded Character Set) et propose une architecture à 4 octets où chaque caractère a une position définie et consiste en 256 plans. Elle permet donc la représentation de caractères occidentaux en les codant sur 32 bits (voir la figure 2).[38]

Nom du code	Nombre de bits	Nombre de caractères
ASCII	7	128
ISO/LATIN1	8	256
UNICODE	16	65536
ISO/IEC10646	32	> 2 milliards

<Figure 2>

3 Les systèmes des grands réseaux de bibliothèques pocédant les Extrême-Orientaux

Un réseau informatisé de bibliothèques consiste à faire travailler en ligne un ensemble de bibliothèques autour d'une base de données bibliographiques communes. Chacun a développé et introduit sa manière ou une fonctionnalité différente.

3-1 OCLC (Online Computer Library Center)

Le premier système **OCLC CJK 350** (dans l'environnement de Windows) est créé en 1986.

Au bout de la génération de **CJK Plus**, la dernière génération CJK PRISM, s'est diffusée dans 45 bibliothèques américaines, y compris la Library of Congress ; la Bibliothèque d'Harvard-Yenching ; la Bibliothèque universitaire de California à Berkeley ; les bibliothèques municipales de Chicago, de Los Angeles et de San Francisco. En plus 5 bibliothèques, en Australie, à Hong Kong, au Singapour, à Taiwan, utilisent aussi le système OCLC CJK[11,12,13], qui permet de

- le catalogage standard en ligne
- imprimer le catalogue sur fichiers et labels au site local
- l'exploitation des bibliographies de la base de données OCLC au système local en caractères CJK ou pas.

Son système de codage est East Asian Character Code (EACC⁷), qui permet d'échanger des enregistrements facilement.[66]

3-2 RLIN (Research Libraries Information Network)

Le RLIN a créé le système RLIN CJK en 1983 qui permet de lire des caractères vernaculaires : chinois, japonais et coréen dans la base de données RLIN. Plus récemment encore RLIN a lancé le projet JACKPHY (Japanese, Arabic, Chinese, Korean, Persian, Hebrew, Yiddish) visant à maîtriser ces alphabets (le cyrillique est traité dès 1986, l'hébreu dès 1988, et l'arabe dès 1990).[2]

Ses participants principaux sont : Université de Yale, Columbia, Michigan, Princeton et Toronto.

Son système de codage Z39.64-1989 (EACC ou East Asian Character Code) est créé par RLG (Research Libraries Group) collaborant avec la Library of Congress et le Chinese Character Analysis Group. La structure de EACC est basée sur le système CCCII⁸ (Chinese Character Code for Information Interchange).[3, 82]

3-3 UTLAS (University of Toronto Library Automated System)

Les données bibliographiques de l'UTLAS provenaient de la Library of Congress, de la Bibliothèque nationale du Canada, de la British Library, de la National Diet Library du Japon, ainsi que des bibliothèques participantes.[35] Il s'est développé au Japon, grâce à la coopération commerciale, par exemple la maison Maruzen qui est une firme spécialisée dans l'édition et dans l'import-export de livre.[2] Son système est capable de gérer les caractères japonais standards (JIS), chinois et coréen. En 1994, UTLAS est racheté par la société canadienne ISM (Information System management).

3-4 NLA (National Library of Australia)

Dès 1996, la NLA a établi le service d'un catalogue en ligne partagé du matériel CJK. Afin d'afficher des caractères CJK dans le système, il est nécessaire d'équiper le logiciel approprié sur un poste de travail, par exemple MASSE⁹, WinMASS, SE JOIGNENT, Eten dans l'environnement de Windows.[16, 19, 21, 22, 26] Mais le système du réseau bibliographique australien, ABN (the Australian Bibliographic Network) n'a pas encore la capacité d'accéder aux caractères CJK.[20] Les utilisateurs d'ABN peuvent accéder

⁷ EACC inclut les caractères idéographiques de "Han", caractères japonais (Katakana et Hiragana) et caractères coréens.

⁸ Ce codage est développé à Taïwan, a des capacités des caractères chinois, japonais et coréen.

⁹ Voir Annexe 3

au système CJK de NLA, mais ils sont limités aux vues des notices romanisées à cause de l'interface.

Le système CJK de la NLA utilise le codage de l'EACC (East Asian Character Code) pour les caractères.[21, 76]

3-5 CNRS (Centre National de la Recherche Scientifique en France)

Depuis 1995, le CNRS à Grenoble (France), a commencé à créer un catalogue collectif chinois de l'Institut d'Extrême-Orient. Dès 1996, il a traité les caractères vernaculaires CJK. Il permet de lire les caractères CJK et d'importer ces mêmes notices. Mais comme pour les autres systèmes, il est nécessaire d'installer une plateforme de la gestion des caractères non-latins pour visualiser directement les caractères CJK.[17] Le CCO (Catalogues Collectifs des Ouvrages) offre la possibilité de télé-charger[70] toutes les fonts (lot de polices de caractères) nécessaires pour afficher correctement les notices utilisant des caractères non-latins.

Pour les japonais, la interrogation se fait en transcription Hepburn¹⁰.

4 Problèmes

Bien que différents, les réseaux présentent beaucoup de similarités. Ils jouent un rôle bibliographique important pour les bibliothèques spécialisées dans les fonds Extrême-Orientaux. Mais pour ce qui concerne l'OCLC, la diffusion du service CJK n'est pas encore très répandue, et l'utilisation du système global est très particulière, elle nécessite par exemple :

- un logiciel approprié
- des méthodes de saisie des caractères particuliers
- une conception propre de données structurées
- une structure des champs

Pour cela, l'échange des notices entre les différents systèmes pose un problème de gestion au niveau de leur récupérations.[5]

¹⁰ Le système de transcription des caractères japonais en roman. (le système Pinyin pour le chinois) Par exemple, Les "e" finaux doivent être prononcés "é", la confusion a donné naissance à Kamikaze (au lieu de "kazé"). Les "s" ne font jamais "z", comme dans asa. Les "h" s'expirent. Le "r" est très proche du "l". Le "u" se prononce entre "ou" et "eu" lorsqu'il est employé dans des mots Japonais, dans les mots chinois le même "u" se prononce plus "u".

5 Conclusion

Le remplacement de l'ASCII par Unicode / IEC 10646 en tant que codage universel des caractères, offre de très grands avantages. Mais, au préalable, un travail de conception des applications est nécessaire. C'est une étape de normalisation fondamentale dans l'implantation de systèmes d'information sur réseau dans le monde qui s'ouvrira probablement sur les possibilités du codage à 32 bits autorisant le codage de termes entiers et de fonctions de commandes informatiques.

Il semble que le système CJK de la NLA (Australie) est plus compatible avec les autres systèmes CJK, parce qu'il est possible d'intégrer les données provenant d'origines diverses : par exemple la Library of Congress, l'OCLC et le RLIN.[3, 31]

Dans tous des cas, il est indispensable de posséder un logiciel approprié pour générer et afficher des caractères CJK sur l'écran.

En outre, la gestion de l'indexation des notices, de l'identification des Kanji (caractères japonais) et des caractères chinois, la création de dictionnaire qui permet de saisir facilement des notices, (par exemple les noms d'auteur) elles ne sont pas facile à normaliser.

Il sera véritablement nécessaire de normaliser tous ces systèmes CJK si l'on veut échanger les données bibliographiques.

Annexe 1

La liste des firmes et des bibliothèques proposant des systèmes de gestion automatisée des bases de données multi-lingues avec possibilités de multiscritp CJK.¹¹

Firme	Ameritech Library services
Pays	Les Etats-Unis
Logiciel principal	UNIX/Pick
Principales séquences type de non-roman couvertes	chinois, japonais, coréen

Bibliothèque	Australian national CJK Sysyem (National Library of Australia)
Pays	L'Australie
E-Mail	lgroom@nla.gov.au
Logiciel principal	INNOPAC, MASSE, JOIGNENT
Principales séquences type de non-roman couvertes	chinois, japonais, coréen

Firme	IME Ltd.
Pays	Le Royaume-Uni
Logiciel principal	TinLib (UNIX et MS-DOS)
Principales séquences type de non-roman couvertes	grec, cyrillien, arabe, thaï, chinois

Firme	International Development Research Center
Pays	Le Canada
Logiciel principal	MINSIS pour MPE/iX, Hp-ux, Solaris, DOS, Réseau local de DOS
Principales séquences type de non-roman couvertes	chinois , arabe, grec, coréen , thaï, devanagiri, cyrillien, bengali, gujerati, punjabi

¹¹ issu de Monica Ertel. A project of the IFLA Section on Information Technology, 1996.

Firme	Innovative Interfaces, Inc.
Pays	Les Etats-Unis
Logiciel principal	UNIX
Principales séquences type de non-roman couvertes	chinois, japonais, coréen, thaï,

Firme	ISM International Canada (autrefois UTLAS).
Pays	Le Canada
Logiciel principal	CATSS (UNIX)
Principales séquences type de non-roman couvertes	chinois, japonais, coréen,

Firme	OCLC
Pays	Les Etats-Unis
Logiciel principal	OCLC avec le passeport (accès à distance d'unité centrale avec la connectivité de PC)
Principales séquences type de non-roman couvertes	chinois, japonais, coréen,

Firme	Research Library Group, Inc.
Pays	Les Etats-Unis
Logiciel principal	RLG (accès à distance d'unité centrale avec le logiciel d'émulation de PC)
Principales séquences type de non-roman couvertes	chinois, japonais, coréen, hébreu, arabe, cyrillien

Annexe 2

Liste des logiciels qui génèrent les caractères CJK.¹²

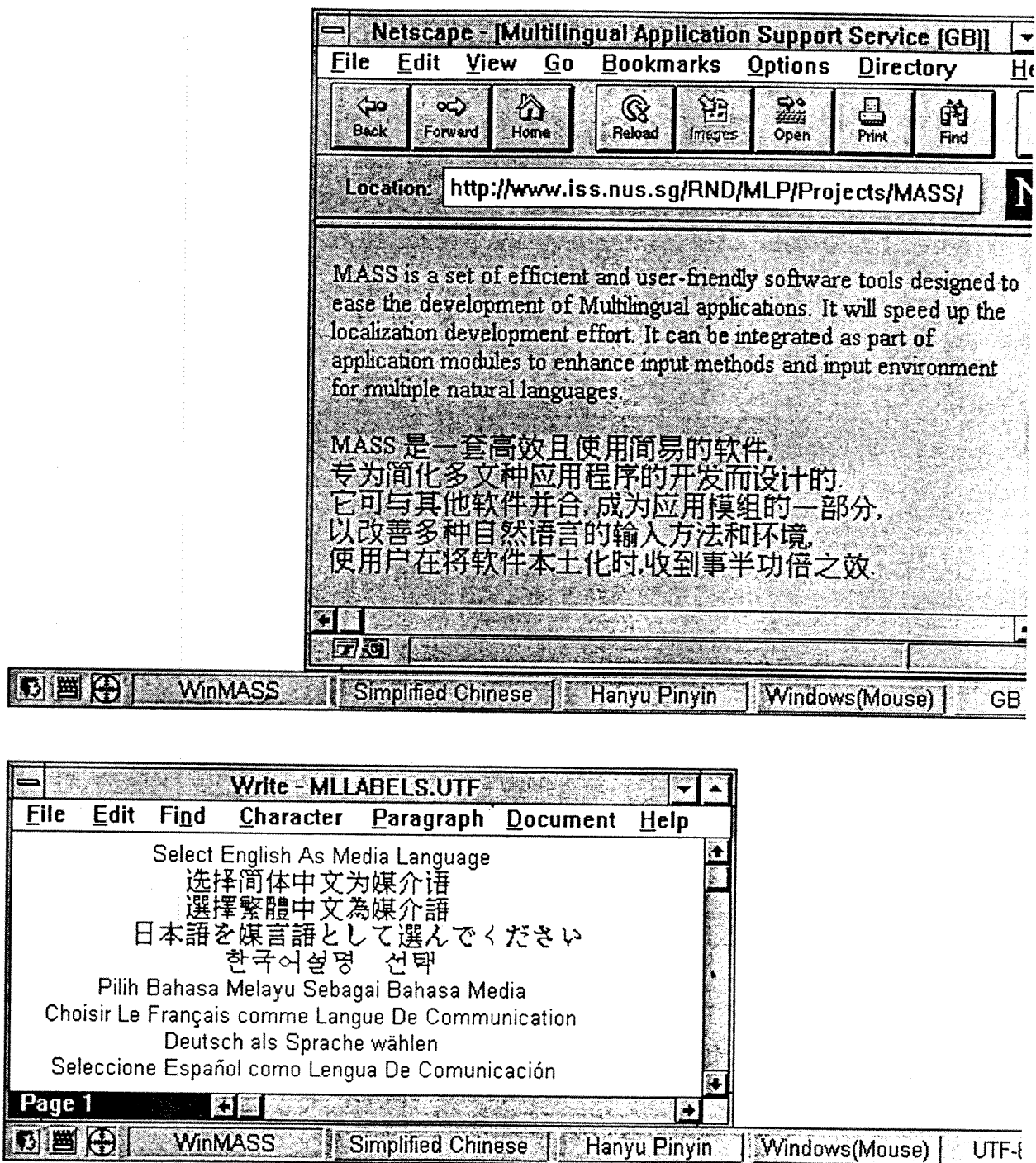
Logiciel	Type	Plateforme	Langues	Fonctions
RichWin	Greffon	Win 3.x/95/NT OS2	CJK	*
Cstar	Greffon	Win 3.x/95	C	*
Njwin	Application	Win 3.x/95	CJK	Lecteur
Njstar	Application	DOS	C	Editique
AsianSuite97	Greffon	Win 3.x/95/NT	CJK	*
TwinBridge	Greffon	Win 3.x/95	C	*
WorldScriptII	Support système	Mac	Codage 2 octets	
Nisus Writer (+WorldScriptII+C /J/K-LK)	Application	Mac	CJK	Editique
Mule	Support système	Unix	CJK	*
Windows*-C/J/K	OS	Windows	CJK	*
FoxPro2-C	Application	Windows	C	SGBD
Lotus AmiPro-C	Application	Windows/OS2	C	Editique
Hamme	Greffon	Win3.x/95	K	*
HWP	Application	Win3.x/95	K	Editique
Eudora-J	Application	Win3.x/95	J	Courriel
MASS	Support système	Unix	CJK	*
WinMASS	Greffon	Win3.x/95	CJK	*
CXTERM	Support système	Unix	C	*
CJK4.1.3	Application	Unix LaTeX	CJK	Editique

* = Fonctions variées des applications compatibles avec la plateforme

¹² issu de Francmonde sur Web

Annexe 3

Exemple du logiciel WinMASS.¹³



¹³ issu de WinMASS Multilingual Application Support Service (MASS) for Microsoft Windows (TM) sur Web

CACIS_HanyuPinyin

ke	1可 2科 3克 4客 5课
科	1科学 2科学家

WinMASS | Simplified Chinese | Hanyu Pinyin | Windows(Keyboard) | GB

CACIS_JianyiBushou

口	一 廿 未	的	一 在 人	阿拉伯	埃及
人	日 一 一	是	会 有 加	哎呀	癌症
八	七 一 十	国	们 十 他	爱好	爱情

WinMASS | Simplified Chinese | Jianyi Bushou | Windows(Mouse) | GB

Annexe 4

Exemple du système CJK Plus par l'OCLC : résultat de recherche par nom d'auteur japonais.¹⁴

OCLC CJK Plus - OLUC Database - Truncated List

Connect Search Edit Action Print Configure Window Help

OLUC Database - Truncated List
Select

OLUC vp 紀田順一/1990-
 Records: 23

Rec#	Name	Title	Publisher	Date	L
1	Kida, Jun'ichir'o,	Bakumatsu Meiji f'uzoku itsuw T'oky'od'o Shu	幕末明治風俗逸話事東京堂出	1993	D
2	Kida, Jun'ichir'o,	Hon no isu /	San'ichi Shob'	1990	D
	紀田順一郎,	本の椅子 /	三一書房,		
3	Kida, Jun'ichir'o,	Hon no j'oh'o jiten /	Shuppan Ny'usu	1991	
	紀田順一郎,	本の情報事典 /	出版ニユ		
4	Kida, Jun'ichir'o,	Izumi Ky'oka.	Chikuma Shob'o	1991	
	紀田順一郎,	泉鏡花.	筑摩書房,		
5	Kida, Jun'ichir'o,	Kindai jibutsu kigen jiten /	T'oky'od'o Shu	1992	
	紀田順一郎,	近代事物起			
6	Kida, Jun'ichir'o,	Konpy'ut'a no	コンピュータ		
	紀田順一郎,	コンピュー			
7	Kida, Jun'ichir'o,	Kosho /			
	ノボリ	ノボリ			

Search OLUC

Query: vp 紀田順一/1990-

Lock Retrieved Record (OCLC# searches only)

Apply default qualifiers

Default qualifiers: /bks

OK Cancel

1 一
2 弍
3 卅
4 市

Input Code: ICHI
HP: CC JH JK
OK
Cancel
NUM: 4

For Help, press F1

¹⁴ issu de Streamlining access to East Asian materials sur Web

Bibliographie

1 OUVRAGES

1-1 Codage des caractères non alphabétiques

[1] Automated systems for access to multilingual and multiscrypt library materials: proceedings of the second IFLA satellite meeting, Madrid, august 18-19, 1993. edited for the Section on Information Technology, the Section on Library Services to Multicultural Populations, and the Section on Cataloguing by Sally McCallum and Monica Ertel. München, New Providence, London, Paris: K. G. Saur, 1994. ISBN 3-598-21797-8.

1-2 Informatisation des bibliothèques

[2] ACQUSSON, Alain. L'informatisation des bibliothèques : historique, stratégie et perspectives. 1995. Paris : Ed. Du Cercle de la Librairie. Collection Bibliothèques. ISBN 2-7654-0604.

2 ARTICLES DE PERIODIQUES ET ARTICLES ELECTRONIQUES TROUVES SUR LE WEB

2-1 Réseau de RLIN

[3] GOSLING, Andrew. National Library of Australia uses RLIN for CJK records. *RLIN Focus*, December, 1996, Issue 23.

<http://rlg.org/r-focus/i232.cjk-oz.html> (interrogé le 20/12/97)

[4] KANEKO, Hideo. RLIN CJK and the East Asian Library community. *Information Technology and Libraries*. 1993, vol.12, no.4, p.423-426.

[5] TSAO, Jai-hsya. The quality and timeliness of Chinese and Japanese monographic records in the RLIN database. *Library Resources & Technical Services*, Jan. 1994, vol.38, no.1, p.60-63.

2-2 Réseau d'OCLC

[6] BOLICK, Hsi-chu. Searching and card production in CJK350 and CJK Plus system. *Committee on East Asian Libraries Bulletin.*, Feb. 1993, no.98, p.22-32.

[7] DEAN, Nita. OCLC CJK TECHPRO service catalogs for university of Southern California.

<http://www.oclc.org/oclc/press/940228.htm> (interrogé le 30/01/98)

[8] KAMEI, Haruko. China and Chinese character (KANJI) to know about Asia 2, CJK systems in the United States OCLC CJK. *Joho no Kagaku to Gijutsu (Journal of Information Science and Technology Association)*. 1997, vol.47, no.7, p.334-339. (en japonais)

[9] KOTAKA, Hisako. CJK users discuss public services and information access at ALA Midwinter. *OCLC Newsletter*, May/June 1997, vol. , p.13-14.

[10] KOTAKA, Hisako. OCLC CJK Plus seminar held in Washington,D.C.. *OCLC Newsletter*, Jul/Aug. 1995, vol.216,p.29-30.

[11] KOTAKA, Hisako. An update on the OCLC CJK Plus System.
<http://www.oclc.org/oclc/cjk/news/214.htm> (interrogé le 20/12/97)

[12] MURPHY, Bob. *et al.* OCLC Asia Pacific Services. *OCLC Newsletter*, Mar/Apr. 1994, vol.208, p.15-27.

[13] OCLC Asia Pacific Service. OCLC CJK Plus System.
<http://www.oclc.org/oclc/ap/prod/cjkplus.htm> (interrogé le 20/12/97)

[14] SPROAT, Eliza. Nanyang Technological University Library uses OCLC CJK Plus system to catalog Chinese collection. *OCLC Newsletter*, Mar/Apr. 1996, vol.220, p.19.

[15] YU, Abraham J. OCLC CJK Plus : a comparison with the CJK350 system. *Committee on East Asian Libraries Bulletin*, Feb. 1993, no.98, p.17-21.

2-3 Catalogage CJK,/ problème de format

[16] CATHRO, Warwick. Current information technology activities at the National Library of Australia. *LASIE*, June 1997, vol.28, no.2, p.6-13.

[17] DELAHAYE, Hubert. CJK open access catalogue.
<http://iias.leidenuniv.nl/iian/iian7/internet/cjk.html> (interrogé le 30/01/98)

[18] DURAND-DASTES, Vincent. Les textes anciens en caractères chinois dans une bibliothèque contemporaine. 1996. Mémoire du DCB de l'ENSSIB.

[19] East Asian Library Resources group of Australia. The Australian National University. *EALRGA Newsletter*, May 1996, no.31.
<http://coombs.anu.edu.au/SpecialProj/NLA/EALRGA/newsletters96/960505.cjknews.html> (interrogé le 20/12/97)

[20] GROOM, Linda. ABN Network Committee : National CJK Project progress report, Feb.1995.
<http://www.nla.gov.au/2/abn/committees/abnnc199519.html>

[21] GROOM, Linda. CJK system being developed by National Library of Australia.
<http://sunsite.nus.sg/bibdb/access/iss10-/iss10-018.html>

- [22] GROOM, Linda. Teamwork on the Internet : the national CJK (Chinese, Japanese, Korean) project. *Provenance*, July 1996, vol.1, no.3.
<http://www.netpac.com/provenance/vol1/no3/features/teamwork.htm>
- [23] HISAMATSU, A. Recent developments in Australian libraries. *Bulletin of Japan Special Libraries Association*. 1993, no.147, p.1-5. (en japonais)
- [24] HOFFMANN, Ellen. Appendix C : library specific computer applications.
<http://www.library.yoku.ca/pubs/compappc.htm> (interrogé le 13/02/98)
- [25] INAHARA, Minoru. The Trends of Library in Asian Countries : CJK service in National Library of Australia. *Gendai no Toshokan*. 1996, vol.34, no.4, p.201-203. (en japonais)
- [26] MULLER, Charles. Web Dictionary of CJK (Chinese-Japanese-Korean) East Asian Buddhist Technical Terms (including text names, personal names, place names, ect.).
<http://ccbs.ntu.edu.tw/acmuller/bdict.htm> (interrogé le 20/12/97)
- [27] NAITO, E ; YAMADA, H. Toward the acceleration of East Asian scholarly information flow : final report of an international joint research aided by the Ministry of Education, Science and Culture of Japan. *International Forum Information and Documentation*, Jan.1993, vol.18, no.1, p.21-26.
- [28] NARITA, Hajime. Language typology and machine translation. *Denshi Joho Tsushin Gakkai Gijutsu Kenkyu Hokoku*. 1996, vol.96, no.158, p.59-66. (en japonais)
- [29] SHEN, Zhijia. CJK : the unique needs of Chinese, Japanese, and Korean language cataloging. *Colorado Libraries*, Winter 1993, vol.19, no.4, p.46-47.
- [30] SMITH-YOSHIMURA, Karen. A decade of Rlin CJK, *RLIN Focus*, August 1993, issue 3.
<http://www.rlg.org/r-focus.html> (interrogé le 20/12/97)
<ftp://ftp.rlg.org/pub/r-focus/3aug93/rlincjk.txt>
- [31] TRAINOR, Julia. National bibliographic access to Asian materials. *3rd National Roundtable on Library & Asia*. February 1995.
<http://www.nla.gov.au/dnc/asiart/TRAINOR.html> (interrogé le 09/02/98)
- [32] WANG, Hsu-Kuang. CJK OPAC at the University of Oregon library. *Committee on East Asian Libraries Bulletin*. October 1993, no.100, p.16-34.
- [33] WITTERN, Christian. Chinese character sets. 1996.
<http://202.232.157.4/iriz/irizhtml/multiling/codes.htm> (interrogé le 09/02/98)

[34] YASUOKA, Koichi. Why is « liang » included in JIS X 0221 ?. *Joho Shori Gakkai Kenkyu Hokoku*. 1997, vol.97, no.80, p.49-54. (en japonais)

2-4 Le réseau Multiscript, multilingues

[35] EILTS, J. Non-Roman script materials in North American libraries : automation and international exchange. *International Cataloguing and Bibliographic Control*, Jul./Sep. 1996, vol.25, no.3, p.51-53.

[36] ERTEL, Monica. A project of the IFLA Section on Information Technology : database of multiscript bibliographic systems, February 1996.

<http://ifla.inist.fr/ifla/VII/s21/1995/sitp1.htm> (interrogé le 13/02/98)

[37] LAZINGER, SS; LEVI, J. Multiple non-roman scripts in Aleph-Israel's research library network. *Library hi tech*. 1996, vol.14, no.1, p.111-116.

2-5 Unicode, ISO/IEC10646

[38] ANDRE, Jacques ; GOOSSENS, Michel. Codage Unicode et ISO/IEC10646. *Cahier GUTenberg*, Mars 1995, no.20,p.1-53.

[39] BARRY, R. K. The role of character sets in library automation : the development of 8 bit sets and Unicode. *International Cataloguing and bibliographic Control*, Jan./Mar. 1997, vol.26, no.1, p.14-17.

[40] BASTE, Sylvie. Recherches sur la normalisation internationale du codage des caractères codes sur plusieurs octets. (Nov. 1997).

<http://hypermedia.univ-paris8.fr/etudiants/baste.htm> (interrogé le 10/02/98)

[41] BETTELS, Jürgen ; BISHOP, E. Avery.Unicode : a universal character code. *Digital technical journal*. Sept. 1993, vol ;5, no.3, p.21-31.

[42] BIGELOW, Charles ; HOLMES, Kris. Création d'une police UNICODE. *Cahier GUTenberg*, Mars 1995, no.20,p.81-102.

[43] FOWLES, Ken. Unicode evolves. *Byte*, March 1997, vol.22, no.3, p.105-110.

[44] HALL, William. S. Internationalization in Windows, Part I : programming with UNICODE. *Microsoft Systems Journal*, Juin 1994, p.58-71.

[45] IRG (Ideographic rapporteur Group). Revised draft of annex : rules of unification and arrangements for CJK unified ideographs. *An unofficial copy of the content of a working group paper of ISO/IEC JTC1/SC2/WG2-IRG meeting*, 1995.

<http://www.stonehand.com/unicode/faq/cjk/annex/annex.html> (interrogé le 20/12/97)

[46] The Unicode Consortium. The Unicode Standard, version1.1. *Unicode Technical Report 4*.

1993.

[47] Unicode le successeur de l'ASCII. *PC EXPERT*. Janvier 1994, p.249-253.

2-6 Logiciel CJK

[48] LIN, Zi-Yu. How to use CJK software to read Chinese, Japanese, and Korean on the Web. *Computers in Libraries*. Nov./Dec. 1997, vol.17, no.10, p.50-54.

3 RAPPORTS DE CONFERENCE

[49] KIM, Jin H. Problems and approaches for Oriental document analysis. *Proceedings of the Fourth International Conference on document analysis and recognition*. 18-20 Aug. 1997, vol.2, p.1111.

[50] NAITO, E. ; AIHARA, N.. Adaptation of a universal standardized subject heading system in the CJK (Chinese/ Japanese and Korean) environment. *62 th IFLA General conference at Beijing*. 25-31 Aug. 1996.

[51] NISHIKIMI, Mikiko ; TAKAHASHI, Naoto, et al. Multilingual text processing system. *Jiko Chino Gakkai Gengo, Onsei Rikai to Taiwa Shori Kenkyukai Shiryo*.1995, vol.11th, p.49-56.
(en japonais)

[52] SUN, Y. Design and implementation of the CJK unified character set on UNIX system. *9th International Unicode conference, Sep. 1996*, no.2.

[53] YAMADA, H. The prospect of CJK technology and future information services in the East Asian Region. *7th Biennial conference Drinking from a fire hose : managing networked information*, Nov.1993, Melbourne (Australie), p.369-385.

4 BREVETS

[54] TAKENOUCI, Hitoshi ; MAEGUCHI, Ryuichi ; KUDOME, Toru. Interactive apparatus. *Omuron Corp*. Brevet no.JP 9584884, 1995-03-15.

[55] YAMAGUCHI, Hiroshige. Method and device for processing information. *Canon Inc*. Brevet no.JP 93296245, 1993-11-26.

5 SITES INTERNET

5-1 Codage des caractères

[56] CJK Codes - Unicode/ISO-10646

<http://www.mit.edu:8001/afs/athena.mit.edu/user/a/K/akbar/Unicode.html>

[57] CJK Code conversion

<http://www.stonehand.com/unicode/faq/cjk/conversion.html#ql>

[58] East Asian Academic Librarians of California

<http://www.lib.berkeley.edu/EAALC/index.html>

[59] East Asian Library of the University of Minnesota

<http://ealib.lib.umn.edu/Ealib.html>

[60] EASL (European Association of Sinological Librarians) - Libraries and East Asian Library Associations

<http://www.uni-kiel.de:8080/ORIENTALISTIK/easl/libra.html>

[61] Francomonde

<http://www.francomonde.org>

[62] Installation des politiques de caractères non-latins, utilisées par le CCO (Catalogues Collectifs des Ouvrages) pour ses bases multi-alphabets

<http://dodge.upmf-grenoble.fr:8001/doc/GDINSTAF.HTML>

[63] Japanese character sets and encoding methods for PCS

http://www.japanese.com.Software.Localization/pc_encoding.html

[64] National Library of Australia : The Chinese Japanese Korean (CJK) Service

<http://www.nla.gov.au/1/asian/ncjk/cjkproj.html>

[65] Notes on CJK Character Codes and Encodings

<http://www.ifess.org/ftp-pub/softwareinfo/cjk-codes/index.html>

[66] OCLC CJK frequently asked questions

http://www.oclc.org/oclc/cjk/cjkwed_faq.htm

[67] Read chinese in net application

<http://www.cathay.net/help/read-chn.shtml>

[68] Subject Index to CDS (Cataloging Distribution Service)

<http://lcweb.loc.gov/cds/sindex.html>

[69] The Unicode Group, Inc.

<Http://www.unicode.com>

5-2 Catalogage CJK

[70] Catalogues Collectifs des ouvrages de la direction du département SHS (Sciences de l'Homme)

et de la Société) du CNRS (Centre National de la Recherche Scientifique)

<http://dodge.upmf-grenoble.fr:8001/catalogues.html>

[71] East Asian Library, University of Minnesota.

<Http://ealib.lib/umn.edu/Ealib.html>

[72] Information about INNOPAC and Innovative Interfaces Inc.

<http://www.iii.com/screens/iiiinfo.htm>

[73] Libraries in the United States with CJK

<http://darkwing.uoregon.edu/~felsing/cela/cjklb.html>

[74] Oriental and India Office Collections, The British Library

<http://www.bl.uk/collections/oriental/>

[75] The National CJK (Chinese Japanese Korean) Service (National Library of Australia)

<http://www.nla.gov.au/asian/ncjk/cjkproj.html>

5-3 Logiciel CJK

[76] Accessing Chinese Character Text on the Internet

<http://darkwing.uoregon.edu/~felsing/cela/hanyu.html>

[77] Accessing Japanese Character Text on the Internet

<http://darkwing.uoregon.edu/~felsing/cela/nihongo.html>

[78] How to Display Japanese Characters on Screen

<http://central.itp.berkeley.edu/~cal/jpcharf.html>

[79] Internet explorer Products Download

<http://www.microsoft.com/msdownload/ieplatform/lang/lang.htm>

[80] Japanese character sets and encoding methods for PCs

http://www.japanese.com/Software.Localization/pc_encoding.html

[81] Reading Korean on the Internet

<http://darkwing.uoregon.edu/~felsing/cela/hangul.html>

[82] RLIN CJK Font Installation Steps for RLIN PC Terminal program for DOS, version E3e

<http://www.rlg.org/jhlcjk.html>

[83] Streamlining access to East Asian materials

<http://www.oclc.org/oclc/cjk/descrip.htm>

[84] WinMASS Multilingual Application Support Service (MASS) for Microsoft Windows (TM)

<http://www.iss.nus.sg:80/RND/MLP/Projects/MASS/winmass/WinMASS.html>

[85] Wombat-Chinese

<http://www.pitt.edu/~ctnst3/chinese.html>

5-4 Téléchargement

[86] Asia Surf

<http://www.crystal-int.com/asurfe.htm>

[87] AsianSuite 97

<http://www.unionway.com>

[88] Asian Viewer

<http://www.twinbridge.com>

[89] Installation des polices de caractères non-latins, utilisées par le CCO pour ses bases multi-alphabets.

<http://dodge.upmf-grenoble.fr:8001/doc/GDINSTAF.html>

[90] NJWIN CJK Multilingual support

<http://www.njstar.com.au/njwin/#1>

[91] Welcome to Download the HOTTEST UnionWay AsianSuite 97 for 60 days Free Evaluation

<http://www.gy.com/uw/demo.htm>