

**E.N.S.S.I.B**  
**ECOLE NATIONALE SUPERIEURE**  
**DES SCIENCES DE L'INFORMATION**  
**ET DES BIBLIOTHEQUES**

**UNIVERSITE**  
**CLAUDE BERNARD**  
**LYON I**

## **DESS en INFORMATIQUE DOCUMENTAIRE**

### **NOTE DE SYNTHÈSE**

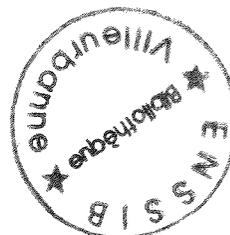
**LES ERREURS DANS LES NOTICES BIBLIOGRAPHIQUES (PAPIER OU  
INFORMATISEES): ERGONOMIE, CONSEQUENCES ...**

**Katarina FANOVA**

**Responsable: M. Jean-Pierre LARDY (URFIST, Lyon)**

1992  
FD  
33

1992



**Titre** : Les Erreurs dans les notices bibliographiques  
(papier ou informatisées): ergonomie,  
conséquences ...

**Nom** : FANOVA Katarina

**Descripteurs français** : Notice bibliographique  
Citation bibliographique  
Erreur, Qualité  
Correction, Détection, Contrôle

### **Résumé**

Etude des erreurs dans les notices bibliographiques; en se basant sur un ensemble des documents obtenus par la recherche automatisée et manuelle, cette note de synthèse dresse une liste de différents types d'erreurs dans les notices bibliographiques et de leurs causes. Ensuite sont traitées les conséquences de ces erreurs et les méthodes pour les détecter, les corriger ou les éviter.

**Key-words** : Bibliographic record  
Bibliographic citation  
Error, Quality  
Correction, Detection, Control

### **Abstract**

A studie of spelling errors in bibliographic records; according to some publications obtained by automatic and manual researches, this synthesis describes different kind of spelling errors in bibliographic records and their causes. Afterward, the consequences and methods in order to detect and correct or to avoid them are given

# TABLE DES MATIERES

<b>INTRODUCTION</b>	4
<b>1 STRATEGIE DE RECHERCHE</b>	5
<b>1.1 Recherche automatisée</b>	5
1.1.1 Interrogation des Bases de Données	5
<i>1.1.1.1 Présentation des Bases de Données interrogées</i>	5
<i>1.1.1.2 Choix des descripteurs</i>	8
<i>1.1.1.3 Equations de Recherche</i>	8
<i>1.1.1.4 Résultats de la Recherche automatisée</i>	9
<b>1.2 Recherche manuelle</b>	9
<b>2. SYNTHESE</b>	10
<b>2.1 Types d'erreurs</b>	10
<b>2.2 Causes d'erreurs</b>	11
<b>2.3 Conséquences d'erreurs</b>	12
2.3.1 Conséquences externes	13
2.3.2 Conséquences internes	13
<b>2.4 Solutions proposées: Détection, correction d'erreurs,         Contrôle qualité</b>	14
<b>Conclusion</b>	17
<b>3. BIBLIOGRAPHIE</b>	18

## INTRODUCTION

Le problème d'erreurs dans les notices bibliographiques touche, probablement sans exception, toutes les Bases de Données et les bibliothèques (que leurs catalogues soient informatisés ou non) et influence ainsi la qualité de leurs services. Une variation de méthodes pour supprimer ces anomalies a été élaborée et testée.

L'objet de ce présent document est de recenser les documents qui traitent le problème d'erreurs dans les notices bibliographiques et d'en faire la synthèse.

Dans une première partie de cette note de synthèse est présentée la stratégie de recherche automatisée et manuelle effectuée.

Une deuxième partie comprend la synthèse des articles recensés. Une liste de différents types d'erreurs et leurs causes est dressée. Notre attention est accordée aux conséquences de ces erreurs et aux solutions proposées par nombreux auteurs.

Les références bibliographiques correspondantes au sujet initial de cette note de synthèse sont classées par ordre alphabétique d'auteurs dans une troisième partie.

# 1 STRATEGIE DE RECHERCHE

## 1.1 Recherche automatisée

La recherche automatisée a été effectuée à l'ENSSIB pendant les TD de recherche documentaire automatisée dans le cadre de la formation de DESSID. Les Bases de Données interrogées étaient: ISA (base n° 202), INSPEC ( base n° 2), LISA (base n° 61), et PASCAL (base n° 144) sur Dialog. Ces Bases ont été choisies en raison de domaine qu'elles couvrent - Science de l'Information. Elles sont donc susceptibles de répondre au sujet de cette note de synthèse.

### 1.1.1 Interrogation des Bases de Données

#### 1.1.1.1 Présentation des Bases de Données interrogées

#### Information science abstracts

**Producteur** : IFI - Plenum Data Company, USA

**Domaines** : **Science de l'information**; Bibliothéconomie, catalogage, classification, édition, enseignement de l'information, indexation, recherche de l'information, reproduction, résumé, science de l'information, technologie des microformes, télécommunication, traduction

**Nature** : références bibliographiques

**Données** : articles (450 périodiques), ouvrages, rapports, compte-rendu de congrès, brevets

**Volume** : 121 000 + 9000 références par an

**Mise à jour** : bimensuelle

**Aides** : Information Science Abstracts, Search manual 1984, User manual 1988

**Serveur** : Dialog (202)

## Inspec

**Producteur** : Institution of electrical engineers (IEE)

**Domaines** : Electronique. Informatique. Physique. Electricité et ingénierie de contrôle: physique atomique et moléculaire; programme, application, systèmes et équipements informatiques; propriétés électriques, magnétiques et optiques; propriétés structurales, thermiques et mécaniques; technologie de contrôles; appareillage électriques et magnétiques; circuits électromagnétiques et optiques; gaz, dynamiques des fluides et plasmas; physique des particules élémentaires; mesures et instruments; **sciences et technologies de l'information et de la communication**; mathématiques; physique nucléaire; systèmes de puissance et applications

**Nature** : références bibliographiques

**Données** : Littérature mondiale: articles de périodiques (80%), actes de congrès (15%), rapports techniques, ouvrages, brevets, thèses universitaires. 2 000 périodiques sont suivis parmi lesquels 200 (dont 7 en français) sont entièrement analysés.

**Début** : 1969

**Volume** : 3 800 000 références + 240 000 références par an

**Mise à jour** : bimensuelle

**Aides** : Magnetic tape services, Inspec user manual (1990), Inspec thesaurus (1991), Inspec classification (1991), Inspec list of journals (1990), LION: logiciel d'apprentissage de la banque

**Publications** : Physic abstracts, Electrical and electronics abstracts. Computer and control abstracts. IT focus. Current papers in physics. Current papers in electrical and electronics engineering. Current papers in computers and control. Key abstracts.

**Serveur** : Dialog (2), BRS, Data-Star, ESA-IRS, ORBIT, STN International

## Library and Information science abstracts

**Producteur** : Library Association Publishing Ltd. Royaume Uni

- Domaines** : **Science de l'information**: automatisation, bibliothéconomie, distribution du livre, édition, reproduction, stockage et recherche de l'information, utilisateurs
- Nature** : références bibliographiques
- Données** : La bibliographie est compilée par la Library Association et par ASLIB (deux organisations anglaises) à partir de tous types de documents: 550 périodiques, ouvrages, communications scientifiques, rapports. Les résumés ne sont disponibles que depuis 1976
- Volume** : 97 000 + 6000 références par an
- Début** : 1969
- Mise à jour** : mensuelle
- Publications** : Library and Information Science Abstracts (6 numéros par an), Current Research in Library and Information Science
- Aides** : Lexique des 3000 descripteurs, Journal in Librarianship and Information Science (1985)
- Serveurs** : Dialog (61), ORBIT (LISA)

## Pascal infodoc

- Producteur** : INIST-CNRS. France
- Domaines** : Sciences et Techniques  
Sciences physiques, Sciences de l'ingénierie, Chimie pure et appliquée, Physique, Chimie, Cristallographie, **Science de l'information**, Technologie, Sciences exactes et appliquées
- Nature** : références bibliographiques
- Données** : Ensemble de banques: Pascal M, Pascal Science de l'Information, Pascal Energie, Pascal Métaux, Pascal Soudage, Pascal Bâtiment, Pascal Biotechnologie, Pascal Médecine Tropicale, Pascal Géode, Pascal Ialine, Pascal Agroline, Pascal Zooline. Analyse de tous les articles de périodiques majeurs français et étrangers ainsi que de rapports scientifiques, thèses, comptes rendu de congrès. Langue d'interrogation: français, anglais et espagnol
- Volume** : 8 000 000 références + 450 000 références par an

**Mise à jour** : mensuelle

**Publications** : Bulletins signalétiques

**Aides** : Profils documentaires, reproduction des documents signalés. Lexique, thésaurus, manuel d'utilisation Pascal

**Serveurs** : Questel (Pascal) 36.29.36.01, IRS-ESA (14), Dialog (144) depuis 1983

### *1.1.1.2 Choix des descripteurs*

Les descripteurs anglais retenus pour la recherche automatisée étaient: **bibliographic record et citation** pour focaliser la recherche sur les notices ou les enregistrements bibliographiques ainsi que sur les citations bibliographiques. Comme le sujet de cette note de synthèse est: les erreurs dans les notices bibliographiques, les descripteurs **error, quality, correction, detection et control** ont été choisis pour préciser le domaine de notre intérêt.

### *1.1.1.3 Equations de recherche*

Les équations de recherche automatisée sont les suivants:

S1 BIBLIOGRAPHIC? (W) (RECORD? OR CITATION? ?)  
S2 S1 AND ERROR?  
S3 (S1 AND QUALITY ) OR S2  
S4 S1 AND (CORRECTION? ? OR DETECTION? ? OR CONTROL? ?)  
S5 S3 OR S4

**Question 1.** Les descripteurs "BIBLIOGRAPHIC RECORD" et "BIBLIOGRAPHIC CITATION" ont été utilisés pour représenter tout ce qui concerne les Notices bibliographiques.

L'opérateur de proximité "(W)" a été introduit pour la recherche de termes adjacents dans l'ordre donné.

La troncature limitée "? ?", dans le cas de terme "CITATION", a été choisie pour tenir compte de "S" au pluriel. Pour les termes "BIBLIOGRAPHIC" et "RECORD" on a jugé plus intéressant d'introduire la troncature illimitée "?". Les termes en question pourraient ainsi prendre successivement la forme de "BIBLIOGRAPHICAL" et "RECORDING".

**Question 2** permet d'obtenir l'intersection de l'ensemble de réponses précédent. La réponse à cette question contient alors les références relatives aux erreurs dans les notices ou citations bibliographiques.

**Question 3.** Les réponses à la dixième question S2 n'étaient pas satisfaisantes. La visualisation des références obtenues par la deuxième équation de recherche a montré que parmi les descripteurs, on pouvait trouver en dehors de "ERROR" aussi "QUALITY" ou "QUALITY CONTROL". Pour élargir la recherche on a donc introduit le terme "QUALITY" qui a été croisé avec la première question S1. On pouvait alors obtenir les références touchant le contrôle qualité des notices bibliographiques.

**Question 4.** L'intersection obtenue par le croisement de la première question S1 et les descripteurs "CORRECTION", "DETECTION" et "CONTROL" avec leur équivalent au pluriel (troncature limitée) donne les références relatives au concept d'erreur.

**Question 5.** L'ensemble final regroupant les questions précédentes.

#### *1.1.1.4 Résultats de la recherche automatisée*

L'interrogation des Bases de Données ISA, INSPEC, LISA et PASCAL a donné les résultats suivants:

- ISA: 7 références, dont 4 pertinentes;
- INSPEC: 6 références, dont 4 pertinentes;
- LISA: 34 références, dont 19 pertinentes;
- PASCAL: 3 références, dont 2 pertinentes.

## **1.2 Recherche manuelle**

La recherche manuelle, dont le but est de compléter la recherche automatisée effectuée précédemment, consistait à examiner la bibliographie des documents obtenus par la recherche précédente afin de repérer les références pertinentes.

Pour localiser les documents primaires qui ont été sélectionnés (dans tous les cas, il s'agissait des revues spécialisées), on a interrogé le Catalogue collectif national des publications en série qui est disponible sur CD-ROM à la bibliothèque de l'ENSSIB.

## 2. Synthèse

Cette note de synthèse basée sur les articles obtenus par la recherche décrite précédemment a pour objet de présenter les différents types d'erreurs, leurs causes, les conséquences que ces causes peuvent entraîner et les solutions proposées par les différents auteurs.

### 2.1 Types d'erreurs

- **erreurs de l'édition** (saisie de données), fautes de frappe: omission, insertion (ceci surtout dans les champs Auteur et Titre), inversion, substitution, transposition, espacement mal placé (un mot divisé en deux ou au contraire, deux mots en un)[23], champ non rempli, inversion des champs [10, 22], erreur de champ [14], erreurs typographiques, "glissement" sur le clavier qui ont souvent des effets dramatiques;
- **fautes d'orthographe**: ditographie ou haplographie (duplication ou omission des lettres qui devraient être répétées[23]);
- **références en double** [3, 4, 6, 19, 31, 32](cas du travail en coopération, de banque de données résultant de la fusion de plusieurs fichiers);
- **erreurs d'interprétation** dues aux lettres non accentuées (marche et marché) ou à la permutation de lettres (comptabilité et compatibilité) [29];
- erreurs dues à la non compréhension** du document par l'indexeur [29];

#### 2.1.1 Autres types d'erreurs

- **erreurs informatiques** de transfert de données entre les sites (problèmes des vieilles bandes magnétiques, des parasites sur la ligne de communication); erreurs de manipulation des opérateurs (au départ et à l'arrivée); perte de support dans le transport ou d'erreur de destinataire; erreur dans le remplissage des formulaires d'accompagnement, sur les étiquettes destinées à la gestion des bandes [29].

## 2.2 Causes d'erreurs

Comme le souligne DYSON[10], il est impossible de quantifier individuellement des causes d'erreurs. Mais la plupart des erreurs sont dues probablement à l'un ou l'autre des quatre facteurs suivants:

- **négligence** qui peut être due à la nature répétitive et fatigante de la saisie. Celle-ci est, sans doute, la première cause d'erreurs;
- **écriture illisible** sur les fiches papier, **conversion et conversion rétrospective** des données en forme lisible par machine;
- **interprétation incorrecte** des instructions;
- "**non-familiarité**" avec les routines des bibliothèques ou de catalogage;
- **la saisie d'ouvrages en langue étrangère** peut être la source de graves erreurs (Selon DYSON[10], les langues moins familières, et particulièrement l'allemand, causent le maximum d'erreurs de saisie);

SUTTER[29], à la différence de l'auteur précédent, parle plutôt de la non-qualité qui affecte le contenu. Il cite d'autres causes d'erreurs:

- **retard dans l'obtention des documents** primaires;
- **retard dans le traitement** des documents: répartition inégale des charges de travail, insuffisance de personnel, documents posant des problèmes particuliers et mis de côté en attendant une décision du chef ou de l'équipe;
- **cas des chaînes résultant de la lecture optique** (problèmes de certaines polices de caractères et de certains corps, de la qualité de l'original où la machine a des difficultés à interpréter deux lettres distinctes mais proches: exemple rn = m);
- **incidences d'évolutions de thésaurus** (perte de liens dus à des non réindexations), insuffisance des relations (absences de synonymes);
- **insuffisante "normalisation" des règles d'écriture** (cas des points dans les sigles ou à la fin de certaines abréviations [18], des traits d'union);
- **incohérence des données** (cas des banques de données numériques);

- **erreurs de cohérence entre les serveurs ou avec le support diffusé** (actualisation selon des rythme ou des modalités différentes, certains sites sont actualisés plus tôt que d'autres);
- **retard dans la mise à jour** (problème de congés et absences dans les petites équipes);
- **insuffisance des données fournies par le producteur** (données non datées) ou non validées, donc de moindre intérêt pour l'utilisateur;
- **incidences perverses des rechargements**: à l'occasion d'amélioration des fonctionnalités et du service aux utilisateurs, un rechargement entraînant un reformatage peut être à double tranchant (certains usagers doivent modifier leur programme local de déchargement); cela implique souvent une réécriture complète de la documentation, ce qui n'est pas toujours fait;
- **le déchargement de données**, le travail en coopération [9]: compatibilité des formats, interprétation locale ou la mauvaise application des normes [8], incohérence des fichiers d'autorité, niveaux différents de catalogage [30]
- **insuffisance ou inadaptation des dispositifs d'assistance** ou de documentation (manuel inexact, non actualisé, incomplet);
- **non remplacement immédiat du personnel partant** (cas du turn-over du personnel), insuffisance de la formation de ce personnel;

### 2.3 conséquences d'erreurs

Les différents types d'erreurs, mentionnées ci-dessus, peuvent avoir de graves conséquences si elles sont présentes dans les notices bibliographiques aussi bien pour les utilisateurs de catalogues en ligne ou de Bases de Données que pour leurs producteurs. On peut alors diviser les conséquences d'erreurs dans les notices bibliographiques en deux parties:

- les conséquences extérieures qui touchent les utilisateurs
- les conséquences intérieures qui touchent les producteurs

### 2.3.1 Les conséquences externes

Une notice bibliographique contient toutes les informations nécessaires au repérage ultérieur du document. Elle lui sert ainsi d'une "carte d'identité" unique et non ambiguë. Le fait, que cette "carte d'identité", qu'elle soit informatisée ou non, comprend les informations erronées, pourrait rendre la recherche du document primaire difficile, sinon impossible. De plus, lors des interrogations de catalogues en ligne ou de Bases de Données, les documents peuvent échapper aux utilisateurs.

### 2.3.2 Les conséquences internes

Pour les producteurs de Bases de Données, les notices bibliographiques représentent les produits qu'ils fournissent à leurs clients - aux utilisateurs. La qualité ou la non-qualité de ces produits va, avec d'autres facteurs, comme la facilité de l'interrogation, par exemple, influencer la fréquence d'utilisation de la Base de Données en question et la satisfaction de la part des utilisateurs. On parle alors de la qualité des Bases de Données (ou bien de la qualité des services intellectuels [21]), qui contribue à leur image de marque et qui joue aussi un rôle primordial sur le marché concurrentiel [29].

Pourquoi ce besoin de qualité? Tout simplement parce que, premièrement, la non-qualité peut coûter les producteurs très cher. Une erreur dans les données fournies par la Prudential Insurance CO., l'omission de trois zéros, a entraîné le banqueroute d'une société et donc a coûté le producteur 11 millions de dollars de dommages et intérêts [23]. Deuxièmement, une étude a prouvé que les utilisateurs de l'OCLC et de la WLN considèrent la qualité de la Base de Données comme la plus importante, et ensuite l'amélioration de service technique [23].

Les erreurs présentes dans les notices bibliographiques causent beaucoup de problèmes dans la gestion interne de Base de Données. Voici quelques exemples:

- **la détection des doublons est plus compliquée;**
- **accroissement considérable du volume de la Base** dû à la présence des doublons (une place de la mémoire ainsi occupée pourrait être utilisée autrement);
- **augmentation de la taille de l'index** causée par des fautes de frappe dans le champ DESCRIPTEURS (exemple: lors de l'interrogation de la Base de Données Pascal, pour avoir le terme "ORGANISATION" en anglais et en français, on a remplacé le "S" par le "?", les résultats étaient: organisation, organization et organiqation).

En général, les conséquences (externes ou internes) d'erreurs dans les notices bibliographiques varient selon la nature d'erreur et selon l'importance du champ où se trouvent les erreurs. Globalement, ce sont les erreurs qui touchent les champs indexables (AUTEUR, TITRE, SUJET, SERIES) qui sont préjudiciables à la qualité. Par ailleurs, une erreur dans le champ TITRE est moins grave qu'une faute dans le champ AUTEUR. Dans le premier cas, chaque mot significatif du titre peut servir d'un moyen d'aide à la recherche. Par contre, dans le second cas, le nom d'auteur n'en est qu'un seul. Le nom erroné peut causer l'échec de la recherche.

De même pour les localisations, classification ou indexation [10]. Une notice qui contient une localisation erronée ne renvoie pas au document primaire. Ceci implique le fait que même si le document est physiquement présent dans le fonds, l'utilisateur ne pourra pas s'en procurer. Les erreurs dans la classification ou l'indexation (les documents mal indexés à cause de la non compréhension du contenu par l'indexeur, absence du contrôle d'autorité dans le cas des notices rapatriées) donnent de fausses résultats de recherche.

#### **2.4 solutions proposés: détection, correction d'erreurs, contrôle qualité**

Les méthodes existantes de détection, correction d'erreurs dans les notices bibliographiques varient en fonction de la nature d'erreur et de systèmes utilisés par les bibliothèques ou les Bases de Données.

- **correcteurs d'orthographe**: de nombreux auteurs [] ont étudié les différents types de fautes d'orthographe commises lors de la saisie d'un document. Ces études ont abouties à des propositions de méthodes de détection d'erreurs et surtout des corrections possibles.

La National Science Foundation (NSF) a lancé une vaste opération d'études de détection et de correction de fautes d'orthographe dans la Base de Données Chemical Abstract Services. Le projet s'appelait SPEEDCOP (SPELLing Error Detection CORrection Project) [25, 26, 33].

- **contrôle d'autorité**: les listes d'autorité contribuent à la qualité de la Base de Données et à la qualité des échanges entre les bibliothèques. Elles assurent un rôle analogue à celui des dictionnaires. Les listes d'autorité exhaustives avec une structure bien définie sont établies avant le démarrage de l'exploitation. Leur établissement exige une préparation soigneusement organisée avant saisie et une correction après saisie.

Le contrôle d'autorité est, normalement, utilisé pour les rubriques qui utilisent les vocabulaires contrôlés (NOM, TITRE, SUJET) [6, 7, 24]. Pour les notices bibliographiques informatisées, le contrôle est effectué lors de chargement de ceux-ci dans le catalogue en ligne. Le fichier d'autorité est contrôlé automatiquement pour vérifier si les rubriques de l'enregistrement sont dans les formes autorisées. Si tel est le cas et s'il n'y a pas d'autres erreurs, la notice bibliographique est entrée dans le fichier maître [14].

- **autocontrôle des données**: certaines données sont structurées de manière à pouvoir en faire des contrôles automatiques. L'utilisation du contrôle de l'ISBN est décrite par l'International ISBN Agency. La méthode est la suivante: on prend l'un après l'autre tous les chiffres du numéro international. On multiplie le premier chiffre par 10, le deuxième par 9, le troisième par 8 et ainsi de suite jusqu'au dernier chiffre. Une fois la multiplication terminée, les résultats intermédiaires sont additionnés entre eux. Pour que l'ISBN puisse être validé, le résultat de l'addition doit être divisible par 11. Cette formule de validation fait partie des sept décrites par ANDERSON ET AL. ECKER & POCH. [23]

- **validation automatique des données** [11] peut être effectuée soit par la vérification du contenu du champ par rapport à son étiquette, soit par la vérification de la cohérence des données contenues dans différents champs dans une notice bibliographique (exemple: le champ identifiant les publications en série restera vide s'il s'agit d'une monographie).

- **contrôle sémantique des données** à entrer, qui offre différentes possibilités d'améliorer la qualité des travaux telles que la saisie unique, les contrôles, les corrections et mises à jour et la gestion des données [16];

- **système interactif de correction d'erreurs**, par exemple: LIAS - Library Acces System - possède une commande (OOPS) qui permet même aux utilisateurs de signaler par un message les erreurs typographiques [28];

- **l'élimination des doublons**: l'utilisation de l'USBC (Universal Standard Bibliographic Code) constitue l'un des moyens adoptés pour la détection et l'élimination des doublons (exemple: le projet QUALCAT - QUALity control in CATaloging est basé sur l'utilisation de l'USBC).

Le problème des doublons est très complexe. Il comprend la définition précise des doublons. La grande importance est accordée à la distinction des imprimés et des éditions d'une monographie. De cette définition dépendront les règles de détection des doublons qui diffèrent en fonction des Bases de Données [3, 4, 31, 32].

Une autre solution pour détecter les doublons est l'utilisation de l'ISBN (International Standard Book Number). L'USBC, comparé avec l'ISBN, prouve un avantage considérable: il est établi en combinant des codes provenant de différents éléments de la notice bibliographique (DATE de publication, TITRE, EDITION, VOLUME). Ainsi, il peut être appliqué pour tous les types de documents (ce qui n'est pas le cas de l'ISBN). De plus, tous les livres ne possèdent pas l'ISBN.

La solution la plus sage reste d'éviter les sources de doublons lors de chargement de notices bibliographiques.

- **édition d'un manuel de contrôle** avec les types d'erreurs et les possibilités de les supprimer [10];

- **formation du personnel:** même si l'informatique offre actuellement les solutions qui facilitent beaucoup la maintenance des catalogues en ligne ou des Bases de Données, il reste toujours à l'homme de s'assurer de sa qualité. Les outils informatiques disponibles allègent les tâches des catalogueurs, mais ne les suppriment pas. Les détections et les corrections automatiques d'erreurs sont de plus en plus performantes, mais pour l'instant, elles ne sont pas totalement fiables [12].

## CONCLUSION

Nous avons vu précédemment que les types d'erreurs rencontrés dans les notices bibliographiques sont très nombreux. Ils proviennent de sources diverses, soit au moment de leur création, soit à cause des traitements qu'on fait subir aux notices bibliographiques (changement de support, par exemple). Une Base de Données dénuée de toutes erreurs est une utopie. Cependant, des procédures de détection et de correction existent, mais elles sont souvent compliquées et ne sont pas toujours fiables.

Les différents types d'erreur mentionnés dans la synthèse sont involontaires, mais il existe aussi des Base de Données qui donnent de fausses informations afin de mieux vendre leurs produits. Quelles dispositions doivent prendre les utilisateurs face à ces problèmes?

Les conséquences d'erreurs peuvent être plus ou moins graves et touchent aussi bien les producteurs de Bases de Données que les utilisateurs. Mieux vaut prévenir que guérir! De nombreuses expériences acquises au cours de l'informatisation des bibliothèques ou lors de la création des Bases de Données montrent qu'il est préférable de prévenir les erreurs et les conséquences qu'elles entraînent en assurant le contrôle permanent et la qualité des travaux que les catalogues en ligne exigent.

# BIBLIOGRAPHIE

Liste alphabétique d'auteurs:

- 1- AUSTER, E., LAWTON, S. B. Search interview techniques and information gain as antecedents of user satisfaction with online bibliographic retrieval. *Journal of the American Society for Information Science*. 1984, vol 35, no. 2, p. 90-130.
- 2- AVRAM, H. Whatever Became of the National Database? *Library Quarterly*. 1983, vol. 53, no. 3, p. 269-278.
- 3- AYRES, F. H., ELLIS, D., HUGGILL, J. A. W., YANNAKOUDAKIS, E. J. The USBC and control of the bibliographic databases. *Information Technology and Libraries*. 1982, vol. 1, no. 1, p. 44-48.
- 4- AYRES, F. H., HUGGILL, J. A. W., YANNAKOUDAKIS, E. J. The Universal Standard Bibliographic Code (USBC): its use for cleaning, merging and controlling large data bases. *Program*. 1988, vol. 22, no. 2, p. 117-132.
- 5- BAKER, A. Computerisation of periodical holdings lists at the University of East Anglia. *UK Serials Group Newsletter*. 1985, vol. 7, no. 1, p. 8-10.
- 6- BILLS, L. G., HELGERSON, L. W. CD-ROM public acces catalogs: database creation and maintenance. *Library Hi Tech*. 1988, vol. 6, no. 1, p. 67-86.
- 7- BURGER, R. H. Artificial Intelligence and Authority Control. *Library Resources & Technical Services*. 1984, vol. 28, no. 4, p. 337-345.
- 8- CHAUVEINC, M. *Le Réseau bibliographique informatisé et l'accès au document*. Paris: Les Editions d'Organisation, 1982. 297 p. Collection Systèmes d'Information et de Documentation. ISBN 2-7081-04756.
- 9- COETZE, H. S., BOON, J. A. Quality control in a bibliographic network. *South African Journal of Library and Information Science*. 1988, vol. 56, no. 4, p. 299-308.
- 10- DYSON, B. Data input standards and computerization at the University of Hull. *Journal of Librarianship*. 1984, vol. 16, no. 4, p. 246-261. DWYER, J. R.

- The Evolutionary Role of Technical Services. *Journal of Library administration*. 1988, vol. 9, no. 1, p. 13-26.
- 11- GABEL, L. Shortcuts I've already found with the PRISM service and PASSPORT software. *OCLC Micro*. 1991, vol. 7, no. 1, p. 8-11.
  - 12- HOLLEY, R. P. The future of catalogers and cataloging. *Journal of Academic Librarianship*. 1981, vol. 7, no. 2, p. 90-93.
  - 13- HORNY, K. Quality work, quality control in technical services. *Journal of Academic Librarianship*. 1985, vol. 11, no. 4, p. 205-210.
  - 14- HUDSON, J. Bibliographic record maintenance in the online environment. *Information Technology and Libraries*. 1984, vol. 3, no. 4, p. 388-393.
  - 15- *Informatique et bibliothèques: Pourquoi et comment informatiser une bibliothèque*. Paris: Cercle de la librairie, 1986. 295 p. Collection Bibliothèques. ISBN 2-7654-0344-9.
  - 16- JONES, B., KASTNER, A. Duplicate Records in the Bibliographic utilities: a Historical review of the Printing versus Edition Problem. *Library Resources & Technical Services*. 1983, vol. 27, no. 2, p. 211-220.
  - 17- JUNEJA, D. C. Quality Control in Data Conversion. *Library Resources & Technical Services*. 1987, vol. 31, no. 2, p. 148-158.
  - 18- KARASICK, A. W. Automated Serials Control: Selection, Acquisition, Utilization. *Serials Librarian*. 1988, vol. 14, no. 3/4, p. 79-87.
  - 19- MACLAURY, K. D. Automatic merging of monographic data bases - use of fixed-length keys derived from title strings. *Journal of Library Automation*. 1979, vol. 12, no. 2, p. 143-155.
  - 20- MATTHEWS, E. W. Quality control for archival tapes. *Law Library Journal*. 1986, vol. 78, no. 4, p. 711-718.
  - 21- MAYER, A. La Qualité des services intellectuels. *International Journal of Services Industry Management*. 1991, no. 1, 13 p.
  - 22- NORTON, N. P. "Dirty Data": A Call for Quality Control. *Online*. 1981, vol. 5, no. 1, p. 40-41.

- 23- O'NEIL, E. T., VIZINE-GOETZ, D. Quality Control in Online Databases. *Annual Review of Information Science and Technology*. 1988, vol. 23, p. 125-156.
- 24- PERSKY, G. The New York Public Library and the Research Libraries Group: an alliance in progress. *The information community: an alliance for progress-proceedings of the 44th ASIS Annual Meeting 1981*. 1981, vol. 18, p. 25-30.
- 25- POLLOCK, J. J., ZAMORA, A. Collection and Characterization of Spelling Errors in Scientific and Scholarly Text. *Journal of the American Society for Information Science*. 1983, vol. 34, no. 1, p. 51-58.
- 26- POLLOCK, J. J., ZAMORA, A. System Design for Detection and Correction of Spelling Errors in Scientific and Scholarly Text. *Journal of the American Society for Information Science*. 1984, vol. 35, no. 2, p. 104-109.
- 27- SOPER, M. E. What you see may not be what you get: errors in online bibliographic records for serials. *Reference Librarian*. 1990, no. 27/28, p. 185-213.
- 28- STRIEDIECK, S. Online Catalog Maintenance: The OOPS Command in LIAS. *Cataloging & Classification Quarterly*. 1985, vol. 6, no. 1, p. 21-32.
- 29- SUTTER, E. *La Qualité des produits et services d'information spécialisée destinés à l'industrie: enjeux et problématique*. 1992. 61 p.
- 30- SUTTER, E. Le Déchargement des données: aspects techniques, économiques et juridiques. *Documentaliste - Science de l'Information*. 1992, vol. 29, no. 3, p. 144.
- 31- YANNAKOUDAKIS, E. J., HUGGILL, J. A. W., AYRES, F. H. *USBC (Universal Standard Book Code): its development and use as method of bibliographic control*. Bradford: University of Bradford, 1984. 102 p. ISBN 0-901945-56-0.
- 32- YANNAKOUDAKIS, E. J., HUGGILL, J. A. W., AYRES, F. H. *USBC (Universal Standard Book Code): its use for union file creation: a feasibility*

*study for a national database*. London: British Library Bibliographic Services Division, 1984. [64] p. ISBN 0-7123-1020-7.

- 33- ZAMORA, E. M., POLLOCK, J. J., ZAMORA, A. The Use of Trigram Analysis for Spelling Error Detection. *Information Processing & Management*. 1981, vol. 17, no. 6, p. 305-316.





\*9590638\*