

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

Rapport de recherche bibliographique

Interrogation des bases de données relationnelles en langage naturel :

- les interfaces ;
- la génération automatique de requêtes SQL à partir d'une question rédigée en langage naturel.

FEDJIDJ Haouari

Sous la direction de

Mr HASSOUN Mohamed
Maître de conférences

E.N.S.S.I.B.
VILLEURBANNE

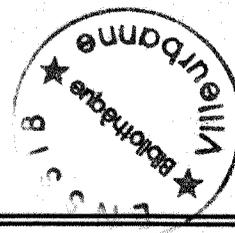
1995

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

Rapport de recherche bibliographique



Interrogation des bases de données relationnelles en langage naturel :

- les interfaces ;
- la génération automatique de requêtes SQL à partir d'une question rédigée en langage naturel.

FEDJIDJ Haouari

Sous la direction de

Mr HASSOUN Mohamed
Maître de conférences

E.N.S.S.I.B.
VILLEURBANNE

1995

ID

24

1995

1

Interrogation des bases de données relationnelles en langage naturel :

- les interfaces ;
- la génération automatique de requêtes SQL à partir d'une question rédigée en langage naturel.

FEDJIDJ Haouari

RESUME :

Le traitement automatique de question en langage naturel est un sujet très vaste, au contour assez flou. Le thème central de notre travail est d'une part de présenter les outils d'interprétation de requêtes adressées en langage naturel à un système de gestion de bases de données relationnelles (SGBD/R) et d'autre part apporter des éléments d'informations sur la génération automatique de requêtes compréhensibles par le SGBD/R. Nous avons essayé dans cette étude, de définir une approche interface caractérisée par une orientation marquée en direction des "non-informaticiens".

DESCRIPTEURS :

Interrogation / Interface / Langage naturel / Bases de données relationnelles / Bases de données.

ABSTRACT :

The automatic treatment of question in natural language is a very vast subject relatively difficult to pin down. The central theme of our work is first of all to present the interpretation tools for queries formulated in natural language to relational databases management system (Relational DBMS) and secondly to bring out elements of information on the automatic generation of understandable queries by Relational DBMS. In this study, we have tried to define an interface approach characterized by a strict orientation towards "non-computer scientist".

KEYWORDS :

Query / Interface / Natural language / Relational database / Relational DBMS / Database.

SOMMAIRE

INTRODUCTION	5
RECHERCHE BIBLIOGRAPHIQUE	7
I Recherche manuelle	8
I . 1 Présentation de la recherche.....	8
I . 2 Bilan de la recherche.....	8
II Recherche automatisée	8
II . 1 Recherche sur CD-ROM.....	8
II . 1 . 1 Choix et description des CD-ROM.....	8
II . 1 . 2 Stratégie de recherche.....	8
II . 1 . 3 Interrogation du CD-ROM "CD-THESES".....	9
II . 1 . 4 Interrogation du CD-ROM "PASCAL".....	9
II . 2 Recherche sur serveurs à partir d'Internet.....	9
II . 2 . 1 Présentation d 'Internet.....	9
II . 2 . 2 Choix et description des bases de données.....	9
II . 2 . 3 Présentation des serveurs.....	10
II . 2 . 4 Présentation des bases de données.....	10
II . 2 . 5 Stratégie de recherche.....	11
II . 2 . 6 Interrogation sur le serveur DIALOG.....	12
II . 2 . 7 Interrogation sur le serveur QUESTEL.....	12
II . 2 . 8 Autres ressources sur INTERNET.....	13
II . 2 . 9 . 1 Interrogation sur le serveur CARL.....	13
II . 2 . 9 . 2 Interrogation sur le serveur GOPHER.....	13
III . 3 Bilan de la recherche.....	13
II . 3 . 1 Analyse des résultats.....	13
II . 3 . 2 Analyse du bruit et du silence.....	14
III Sélection et obtention des documents	15
IV Conclusion de la recherche bibliographique	15
SYNTHESE	16
I Introduction	17
I . 1 Préambule.....	17
I . 2 Notions de linguistique.....	17
I . 3 Présentation générale d'une interface.....	18
I . 4 Langage procédurale et langage non-procédurale.....	18
I . 5 Classification des langages de requêtes.....	18

II Développement	19
II . 1 Introduction de la notion d' interfaces "non-procédurales"	19
II . 2 Présentation des interfaces pour "non-informaticiens"	19
II . 2 . 1 Les interfaces en langage naturel.....	19
II . 2 . 1 . 1 Introduction.....	19
II . 2 . 1 . 2 Généralités sur les systèmes en langage naturel.....	20
II . 2 . 1 . 3 Modèle d'architecture.....	20
II . 2 . 1 . 4 Exemples d'architectures.....	21
II . 2 . 1 . 5 Les références obtenues	21
II . 2 . 2 Présentation sommaire d'autres types d'interfaces.....	22
II . 2 . 2 . 1 Les interfaces graphiques.....	22
II . 2 . 2 . 1 . 1 Introduction.....	22
II . 2 . 2 . 1 . 2 Les références obtenues	22
II . 2 . 2 . 2 Les interfaces relations universelles.....	22
II . 2 . 2 . 2 . 1 Introduction.....	22
II . 2 . 2 . 2 . 2 Les références obtenues	22
II . 3 La génération de requêtes à partir du langage naturel.....	22
II . 3 . 1 Introduction de la notion de paraphrase	23
II . 3 . 2 Exemple de la paraphrase sur SAPHIR	23
II . 3 . 3 Les références obtenues.....	24
 III Conclusion	 24
 BIBLIOGRAPHIE	 26
I Format des notices	27
II Classement par type de références	27
III Présentation de la bibliographie	28
III . 1 Monographies.....	28
III . 2 Rapports de recherches.....	28
III . 3 Thèses	28
III . 4 Articles de périodiques.....	29
III . 5 Actes de congrès.....	33

INTRODUCTION

Pour l'interrogation des bases de données, l'idéal serait de mener un dialogue oral, en langage naturel avec la machine. Les problèmes rencontrés pour reconnaître de façon automatique la parole en continue, rendent difficile la réalisation d'un système opérationnel. La communication "homme-machine" doit donc passer par le biais d'un clavier.

Les performances des systèmes informatiques ne cessent de croître et permettent d'envisager des communications "homme-machine" de plus en plus proche du langage naturel. Pour ce faire, on aimerait que ce dernier devienne l'outil courant de communication avec l'ordinateur. L'utilisation du langage naturel comme moyen de communication entre l'homme et la machine requiert en premier lieu la compréhension du message formulé dans la langue de l'utilisateur, ensuite l'élaboration d'une réponse, enfin une formulation de cette réponse dans la langue de l'utilisateur.

Notre synthèse consistera donc à donner un aperçu global sur le sujet de l'interrogation des bases de données relationnelles en langage naturel. Ce travail sera présenté à partir d'une sélection des documents les plus pertinents, issus d'une bibliographie orientée sur une description des interfaces en langage naturel essentiellement, et sur les produits de génération de requêtes au format SQL.

Enfin, cette étude présente les interfaces en langage naturel existantes sur le marché national et international, mais n'identifie pas les fournisseurs, ni les coûts de ces produits.

RECHERCHE BIBLIOGRAPHIQUE

I RECHERCHE MANUELLE

I . 1 Présentation de la recherche

La recherche manuelle n'avait pas pour but d'être exhaustive, mais de permettre une première approche du sujet. Elle a fourni un premier aperçu de la littérature concernant le sujet (principalement américaine et anglaise). Enfin, elle a servi de préambule à la recherche automatisée en fournissant une première idée du vocabulaire à utiliser.

Cette recherche s'est surtout orientée au dépouillement de revues informatiques (Online, Nouvelles Technologies de l'Information, 01 Informatique, Le Monde Informatique) ; ainsi que la lecture d'ouvrages, se rapprochant du sujet au travers des fonds documentaires des principales bibliothèques de Lyon. Cette consultation a été complétée par l'analyse de notes de synthèses des promotions précédentes, et par un recueil d'articles remis par le demandeur.

I . 2 Bilan de la recherche

Cette approche manuelle, nous a permis d'accéder à certaines sources d'informations primaires, et de recenser six références pertinentes.

Les références concernant l'interface SAPHIR ont été intégrées dans la bibliographie. Cette interface a été une des premières à utiliser les techniques de l'intelligence artificielle.

II RECHERCHE AUTOMATISEE

II . 1 Recherche sur CD-ROM

II . 1 . 1 Choix et description des CD-ROM

Les CD-ROM "CD-THESES" et "PASCAL" ont été sélectionnés pour leur spécificité, afin de permettre de compléter la recherche manuelle.

CD-THESES :

Produit de la base de données TELETHESE sur CD-ROM. Cette base regroupe toutes les thèses de doctorats et de diplômes de docteurs ingénieurs. Fin 1993, plus de 245000 références.

PASCAL :

Produit à partir de la banque de données PASCAL. Ce CD-ROM recense environ 450000 références dans les domaines de la science, technologie et médicale.

II . 1 . 2 Stratégie de recherche

A la différence de CD-THESES, les interrogations sur PASCAL ont été opérées en anglais, car cette langue offre l'avantage d'exprimer en un mot une notion qui en requiert plus en français (DATABA* ramène par exemple les termes database et databank).

Sur les deux CD-ROM, nous n'avons pu utiliser des opérateurs de proximités, ce qui a alourdi les requêtes. Cette recherche a été assez large sans limitation de date et de langue.

II. 1. 3 Interrogation du CD-ROM "CD-THESES"

Utilisation de l'édition de 1992 du CD-ROM "CD-THESES" :

Interrogation N°1 :

SUJ = "LANGAGE NATUREL" ET SUJ = "INTERROGATION BASE DONNEE"

Résultat : Deux références pertinentes dont une dans DIALOG / PASCAL.

Interrogation N°2 :

SUJ = "LANGAGE NATUREL"

Résultat : Deux références pertinentes.

II. 1. 4 Interrogation du CD-ROM "PASCAL"

Utilisation des éditions suivantes du CD-ROM "PASCAL" :

Janvier / Juin 1994 et Juillet / Septembre 1994

Janvier / Septembre 1992 et Octobre / Décembre 1992

Janvier / Septembre 1991 et Octobre / Décembre 1991

Janvier / Décembre 1990

Interrogation N°1 :

DXA = "NATURAL LANGUAGE" ET DXA = DATABASE

Résultat : Sept références pertinentes dont une dans DIALOG / INSPEC.

Interrogation N°2 :

DXA = "RELATIONAL DATABASE"

Résultat : Deux références pertinentes.

Le mode expert a été utilisé, ce qui nous a permis des interrogations plus ciblées.

II. 2 Recherche sur serveurs à partir d'Internet

II. 2. 1 Présentation d'Internet

Internet est un terme générique qui désigne la multitude de réseaux utilisant les protocoles TCP / IP (Transmission Control Protocol / Internet Protocol) dans le monde et dont NSFnet (National Science Foundation Network) constitue le coeur.

Internet est un réseau, présent aujourd'hui dans plus d'une centaine de pays, au sein desquels le nombre de serveurs double chaque année. Il regroupe plus de 40000 réseaux et 30000000 d'utilisateurs dans le monde en 1994.

II. 2. 2 Choix et description des bases de données

Après consultation du Répertoire des banques de données professionnelles 1993 édité par l'ADBS (Association française des Documentalistes et des Bibliothécaires Spécialisés) et parmi les bases de données accessibles, nous avons sélectionné pour leur domaine de couverture du sujet, les bases de données suivantes :

Les bases les plus intéressantes sur le serveur DIALOG :

- COMPENDEX-PLUS.....N° 8
- COMPUTER-DATABASE.....N° 275
- INSPEC.....N° 4
- PASCAL.....N° 144

Sur le serveur QUESTEL :

- FRANCIS.....domaine "linguistique".

Les bases secondaires sur le serveur DIALOG :

- ISA...: Sciences de l'information.....N° 202
- LISA...: Sciences exactes et appliquées.....N° 61
- ERIC...: Sciences de l'information et linguiste.....N° 1

Remarque : Ces trois dernières bases n'ont pas été sélectionnées, après une première phase d'interrogation.

II . 2 . 3 Présentation des serveurs

DIALOG :

Serveur américain couvrant la plupart des domaines. Il regroupe environ 500 bases de données.

QUESTEL :

Serveur français orienté dans les domaines :

- sciences et techniques ;
- brevets et marques ;
- informations sur les entreprises.

Ce serveur regroupe environ 180 bases de données.

II . 2 . 4 Présentation des bases de données

COMPENDEX-PLUS :

* Domaine : Base couvrant les domaines de l'ingénierie ainsi que les sciences et techniques (1970).

* Producteur : Engineering information (EI). USA.

* Volumes : 2600000 références dont 160000 par an. Environ 2500 périodiques, publications d'organisations, sociétés d'ingénierie, actes de congrès, monographies et rapports.

COMPUTER-DATABASE :

- * Domaine : base orientée essentiellement en informatique (1983).
- * Producteur : Information access company (IAC). USA.
- * Volumes : 192000 références dont 50000 par an. Des articles de plus de 180 périodiques en langue anglaise ainsi que rapports, livres et newsletters.

INSPEC :

- * Domaine : C'est une importante base bibliographique sur l'électronique, l'informatique et la physique (1969).
- * Producteur : Institution of electrical engineers (IEE). Grande-Bretagne.
- * Volumes : 3800000 références dont 240000 par an. Les données couvrent la littérature mondiale et comprennent en grande partie des articles de périodiques et des actes de congrès.

PASCAL :

- * Domaine : Base multidisciplinaire et sous-bases sectorisées (1973).
- * Producteur : Centre national de recherche scientifique-Institut national pour l'information scientifique et technique (CNRS-INIST). France.
- * Volumes : 9145410 références dont 600000 par an. Les données proviennent d'articles de périodiques, de congrès, de thèses, de rapports de divers pays.

FRANCIS :

- * Domaine : Base couvrant les sciences humaines, sciences sociales et sciences économiques (1972).
- * Producteur : CNRS-INIST. France.
- * Volumes : 1400000 références dont 80000 par an. Les données issues d'articles de périodiques, d'ouvrages, de rapports et de comptes-rendus de congrès.

II . 2 . 5 Stratégie de recherche

- La recherche des mots clés a été validée à partir du thésaurus : "The Original ROGET'S" Thesaurus of English words phrases.
- Le mot clé Online n'a pas été utilisé, car celui-ci est trop large par rapport au type d'interrogation que l'on souhaite réaliser. Générant trop de bruits, ce mot a été écarté.
- De même, le mot clé SQL (Structured Query Language) n'a pas été utilisé, afin de ne pas orienter la recherche uniquement sur :
 - . les bases de données relationnelles n'utilisant que ce langage de requête ;
 - . la génération de requêtes n'utilisant que le modèle SQL.
- A partir du Basic Index, on pourra rechercher les termes pour construire les équations sur les champs : Titres, Résumés, Descripteurs et Termes Génériques.
- L'utilisation du catalogue DIALENDEX n'a pas été nécessaire pour repérer les bases de données.
- Pour les bases COMPENDEX-PLUS / INSPEC / PASCAL, on appliquera la procédure "ONE SEARCH".
- La base COMPUTER-DATABASE sera interrogée séparément, celle ci n'acceptant pas le critère de sélection sur la langue de publication du document.
- Au vu des résultats des premières interrogations, on limitera la recherche sur deux critères :

. la date sera fixée à partir de 1987, limite raisonnable par rapport à l'avancement des recherches dans ce domaine.

. les documents en langue : Japonnaise, Chinoise et Coréenne ont été écartés.

- On a utilisé pour l'élaboation des équations de recherches les outils suivants :

	DIALOG	QUESTEL	SIGNIFICATION
Opérateurs logiques	OR	pas utilisé	ou inclusif
	AND	ET	et
	NOT	pas utilisé	sauf
Opér. de proximités	W	AV	un terme au plus entre les deux termes
Opér. de troncatures	??	?	0 ou 1 caractère vers la droite

II . 2 . 6 Interrogation sur le serveur DIALOG

Interrogation N°1 :

B 4,8,144

S1 (QUERY OR QUERIES) AND (INTERFACE? ?) (ref=2899)

S2 NATURAL LANGUAGE (ref=13075)

S3 (RELATIONAL(W)DATABASE? ? OR RELATIONAL(W)DBMS) (ref=12212)

S4 S1 AND S2 AND S3 (ref=91)

S5 S4/1987:1995 (ref=62)

S6 S5 NOT (LA=JAPANESE OR LA=CHINESE OR LA=KOREAN) (ref=55)

S7 RD S6 (ref=47)

Interrogation N°2 :

B 275

S1 (QUERY OR QUERIES) AND (INTERFACE? ?) (ref=2175)

S2 NATURAL LANGUAGE (ref=1291)

S3 (RELATIONAL(W)DATABASE? ? OR RELATIONAL(W)DBMS) (ref=5128)

S4 S1 AND S2 AND S3 (ref=254)

S5 S4/1987:1995 (ref=21)

II . 2 . 7 Interrogation sur le serveur QUESTEL

Interrogation :

1 INDEX/DO LINGUISTIQUE (ref=79809)

2 LANGAGE AV NATUREL (ref=231)

3 1 ET 2 (ref=58)

4 BASE? AV RELATIONNELLE (ref=28)

5 2 ET SGBD (ref=2)

6 2 ET INTERFACE? (ref=9)

7 5 OU 6 (ref=10)

II . 2 . 8 Autres ressources sur INTERNET

II . 2 . 8 . 1 Interrogation sur le serveur CARL

Serveur CARL / Base de données UNCOVER :

Uncover est une base de données bibliographiques qui recense 14000 journaux dans les domaines des sciences, techniques et biomédicales. Cette base est consultée sur le serveur CARL à partir du logiciel TELNET à l'adresse PAC.CARL.ORG ou 192.54.81.128.

Interrogation N°1 :

QUERY INTERFACE NATURAL LANGUAGE RELATIONAL DATABASE

Résultat : Deux références pertinentes dont une dans DIALOG / INSPEC.

Interrogation N°2 :

QUERIES INTERFACES NATURAL LANGUAGE RELATIONAL DATABASE

Résultat : Une référence pertinente.

Interrogation N°3 :

QUERY RELATIONAL DATABASE

Résultat : Sept références pertinentes.

Interrogation N°4 :

QUERIES RELATIONAL DATABASE

Résultat : Trois références pertinentes.

II . 2 . 8 . 2 Interrogation sur le serveur GOPHER

Serveur GOPHER / VERONICA :

VERONICA est un outil de recherche pour naviguer à travers les serveurs GOPHER. Cet outil permet de saisir des critères de recherche, composés de un ou plusieurs mots-clés pour les documents désirés.

Chemin suivi : choisir *Autres serveurs d'information*
puis *Serveurs Gopher en France*

puis *Université de Nice*

puis *Serveurs Gopher et recherche dans VERONICA*

puis *Les serveurs Gopher Nord Américain*

Interrogation :

INTERFACE NATURAL LANGUAGE

Résultat : Une référence pertinente.

II . 3 Bilan de la recherche

II . 3 . 1 Analyse des résultats

Nos expériences sur CD-ROM furent longues car elles ont demandé un temps d'adaptation aux méthodes et aux procédures de recherche. Ces travaux ont toutefois été vécus comme une opportunité de s'exercer sans souci de coût. Les premières recherches ont servi de préparations efficaces à l'interrogation en lignes des bases de données.

Analyse de la recherche sur DIALOG / QUESTEL :

	Nombre de références obtenues	Nombre de références pertinentes	Taux de pertinence %
INSPEC	42	33	78,6
COMPENDEX-PLUS	1	1	100
PASCAL	4	3	75
COMPUTERDATABASE	21	14	66,7
FRANCIS	8	3	37,5

Ce tableau comparatif montre un taux de pertinence satisfaisant pour les bases interrogées sur DIALOG, notamment la base INSPEC qui a été la plus significative en proposant 33 références pertinentes sur un total de 78 pour la recherche automatisée. Sur le serveur QUESTEL, ce résultat très moyen peut s'expliquer par au moins deux raisons :

- le domaine de couverture de FRANCIS, pouvait apporter des éléments d'informations complémentaires par sa spécificité dans le traitement du langage ;
- l'interrogation à partir de descripteurs en français a généré un faible nombre de références, phénomène qui aurait pu s'inverser par une interrogation en anglais (cf II . 1 . 2).

Une interrogation particulière de COMPENDEX-PLUS a été réalisée à partir du thésaurus d'EI-Compendex, pour la recherche de descripteurs. Cette interrogation a généré 15 références non satisfaisantes, ci-dessous l'équation utilisée :

```
S1 QUERY(W)LANGUAGES
S2 USER(W)INTERFACES
S3 (RELATIONAL(W)DATABASE(W)SYSTEMS OR RELATIONAL(W)DBMS)
S4 S1 AND S2 AND S3
S5 S4/1987:1995
```

Pour les autres ressources Internet, une redéfinition des requêtes plus ou moins larges, a permis d'identifier des références complémentaires, qui figuraient dans des revues non proposées par les autres serveurs.

II . 3 . 2 Analyse du bruit et du silence

Cette étude comparative a été la plus représentative sur le serveur DIALOG. La prise en compte des descripteurs obtenus à partir du Basic Index tels (QUERY OR QUERIES), (RELATIONAL DBMS) a permis d'éliminer beaucoup de bruit et d'élargir le nombre de références. Ce résultat a eu pour effet de réduire le silence et de passer de 31 à 54 références.

Les interrogations avec un croisement des descripteurs de ces deux façons :

- NATURAL(W)LANGUAGE(W)INTERFACE? ?
- (NATURAL(W)LANGUAGE? ?) AND (INTERFACE? ?)

ont eu le mérite de mettre en avant la mauvaise structuration des questions.

Pour arriver à proposer une interrogation relativement satisfaisante sur le serveur DIALOG, il nous a fallu établir différents niveaux de pondérations :

- * niveau 1 : - la capacité d'*interrogation* et les produits d'*interfaces* ;
 - à partir du *langage naturel* ;
 - appliqué aux *SGBD/R*.
- * niveau 2 : - la combinaison de ces trois composantes pour créer l'environnement nécessaire pour cibler le sujet.
- * niveau 3 : - la limitation sur la langue et la date de publication.

III SELECTION ET OBTENTION DES DOCUMENTS

La sélection des articles a été assez aisée dans les bases des serveurs DIALOG / QUESTEL, mais plus difficile et aléatoire sur le serveur CARL. De même les résumés étaient courts voire inexistant dans la base UNCOVER.

L'obtention des documents les plus significatifs, a été réalisée suivant un schéma classique :

- localisation des documents à l'aide :
 - . du CD-ROM "MYRIADE" (Périodiques / Congrès) ;
 - . du logiciel LORIS / DORIS (Ouvrages de l'ENSSIB) ;
 - . du logiciel de la Bibliothèque Universitaire de Lyon I (Thèses Françaises) ;
- demande de Prêt Entre Bibliothèques pour les Thèses ;
- consultation de microfiches.

IV CONCLUSION DE LA RECHERCHE BIBLIOGRAPHIQUE

L'analyse des résultats de la recherche bibliographique montre d'une part que le sujet n'est pas récent, et d'autre part que la littérature qui la concerne est en très forte proportion anglaise. Une majorité de références se trouvent dans les articles de périodiques et les actes de congrès, ceci peut s'expliquer par l'évolution constante de ce sujet et fait encore aujourd'hui l'objet d'actives recherches. Les documents sélectionnés traitent relativement bien le domaine de l'interrogation des bases de données relationnelles.

En résumé ce travail de recherche bibliographique de part le temps passé, les réflexions qu'il a engendré et les contacts personnels qu'il a imposé, nous a permis de réaliser une recherche consciencieuse. Cette étude a eu le mérite de nous confronter aux problèmes liés à la recherche documentaire, qu'elle soit manuelle ou automatisée. Enfin, nous avons pu mettre à jour nos connaissances, et nous l'espérons celles du demandeur concernant l'interrogation des bases de données relationnelles en langage naturel.

I INTRODUCTION

I. 1 Préambule

La construction d'une interface utilisateur est un problème d'une grande complexité dont les tenants et les aboutissants sont loin d'avoir été complètement dégagés. Cela peut s'expliquer par la diversité :

- des utilisateurs, ce qui entraîne une inflation de contraintes ergonomiques ;
- des applications potentielles et des objets manipulés ;
- des environnements de travail, matériels de base sur lesquels devra pouvoir fonctionner l'interface ;

mais aussi par le dialogue "homme-machine" lié à l'activité mentale de l'homme, illimitée et non canalisée par un programme.

Nous profitons de ce préambule pour rappeler la définition d'un SGBD/R :

Un Système de Gestion de Base de Données Relationnelles est un ensemble de logiciels permettant d'offrir aux utilisateurs les moyens de stocker, d'accéder et de mettre à jour des données, tout en assurant leur sécurité. Il est caractérisé par :

- le choix du modèle relationnel pour les niveaux logiques du SGBD ;
- le choix d'un langage complet pour l'interface utilisateur.

Nous rappellerons aussi que, SQL est un langage de définition et de manipulation de données obéissant à un schéma dit "relationnel" permettant d'enregistrer, modifier, et supprimer des données dans la base et de formuler enfin des interrogations dans cette même base.

I. 2 Notions de linguistique

En linguistique¹, on considère cinq niveaux dans la langue :

- *morphologique*...: reconnaissance du mot sous différentes formes par exemple conjugaisons, déclinaisons ;
- *lexicale*.....: correspondance du mot une fois reconnu avec les informations dont on dispose sur ce mot ;
- *syntaxique*.....: agencement du mot dans une phrase ;
- *sémantique*.....: correspondance des situations du monde réel aux structures reconnues par le niveau syntaxique ;
- *pragmatique*.....: interprétation de ces situations dans le contexte plus général d'un échange d'information entre l'auteur et le lecteur.

¹Dans cette analyse, on ne détaillera pas le traitement de la langue naturelle, mais on pourra se reporter à la note de synthèse de CHALON, J.M. / ID 7 / 1991 qui traite de ce sujet.

I . 3 Présentation générale d'une interface

Une interface en langage naturel est un système plus ou moins complexe ayant pour tâche de traduire des expressions (questions, ordres, etc...) d'une langue naturelle (française, anglaise, etc...) dans celle d'un langage formel particulier.

Par exemple :

- quels sont les départements de plus de 2 millions d'habitants ?
- traduction dans un SGBD/R :

```
SELECT département, population
FROM table 1
WHERE population > 2000000
```

L'utilisateur est dispensé d'apprendre un langage formel ; le langage utilisé est un sous-ensemble plus ou moins riche de sa langue.

I . 4 Langage procédurale et langage non-procédurale

L'utilisateur n'a pas à connaître l'organisation et la structure des données pour formuler ses requêtes ou ses ordres. C'est par exemple toute la différence entre un langage de requête de type SGBD et un langage naturel :

- le premier est *procédural* :
 - . l'utilisateur doit énumérer dans un certain ordre l'ensemble des opérations à effectuer pour parvenir à l'information recherchée.
- le second est *non-procédural* ou *déclaratif* :
 - . indépendance de toute forme de stockage des informations recherchées.

I . 5 Classification des langages de requêtes

Dans les bases de données relationnelles, les langages de requêtes se caractérisent par leur complétude², puissance, et simplicité. Ces langages peuvent être regroupés en deux classes principales correspondant à la dualité de définition d'une relation, qui peut être vue :

- * (1) soit comme un *ensemble* :
 - la classe correspondante est celle des *langages ensemblistes*.
 - les langages ensemblistes n'utilisent pas de variables, une requête y décrit le lien ensembliste entre les attributs résultants, un ensemble défini de qualifications et inter-relations. Ces requêtes sont traduites par une série d'inclusions ou d'égalités ensemblistes.

Parmi les langages ensemblistes figurent :

- . les langages algébriques
- . les langages SEQUEL, SQL,...

²Complétude : propriété d'une théorie déductive consistante où toute formule est décidable.

* (2) soit comme un *prédictat* :

la classe correspondante est celle des *langages prédicatifs*.

les langages prédicatifs peuvent être divisés en deux sous-classes suivant que la relation s'applique :

(2.1) sur les tuples :

. la sous-classe correspondante est celle des langages relationnels orientés tuples ; langages qui se rapprochent du langage naturel.

(2.2) sur les domaines :

. la sous-classe correspondante est celle des langages relationnels orientés domaines, langages qui facilitent la réalisation d'interfaces conviviales.

II DEVELOPPEMENT

II . 1 Introduction de la notion d'interfaces "non-procédurales"

Les interfaces actuelles entre un utilisateur et une base de données relationnelles se limitent généralement, à des langages non procéduraux. L'élargissement du concept d'interface est un problème ouvert dans le domaine de la recherche. Après avoir décrit les modèles de base qui serviront de cadre formel à la plupart des interfaces non procédurales, nous présentons ici les fonctionnalités des interfaces pour non-informaticiens.

II . 2 Présentation des interfaces pour "non-informaticiens"

Dans cette analyse, nous nous intéresserons à trois types d'interfaces :

- principalement aux interfaces en langage naturel³ ;
- secondairement :
 - . aux interfaces graphiques ;
 - . aux interfaces basées sur la notion de relation universelle.

II . 2 . 1 Les interfaces en langage naturel

II . 2 . 1 . 1 Introduction

Avant de présenter quelques systèmes offrant la possibilité à l'utilisateur d'interroger une base de données en langage naturel, nous commençons cette description des interfaces par une introduction visant à mettre en évidence les raisons du développement de ce type d'interface textuelle, mais également leurs limites.

Cette introduction servira aussi de tremplin pour présenter des interfaces en langage naturel pour le domaine documentaire :

- les interfaces qui offrent une aide en ligne pour la recherche en mode conversationnel : "front-end" ;
- les interfaces opérant comme passerelles entre l'utilisateur et les bases de données auxquelles on souhaite accéder : "gateway".

³Le langage naturel inclu aussi le langage libre qui se limite plus à un contexte donné.

II . 2 . 1 . 2 Généralités sur les systèmes en langage naturel

Les interfaces utilisateurs fournies par les systèmes de SGBD permettent de manipuler la base par le biais de langage d'interrogation. L'utilisation du langage naturel, pour interroger les bases de données, est une idée très attrayante ; les systèmes en langage naturel correspondent aux besoins de l'utilisateur non-informaticien, puisque le langage naturel ne demande ni apprentissage, ni connaissance particulière.

L'ambiguïté du langage naturel et les difficultés de construire des systèmes complets représentent l'inconvénient majeur de ces systèmes :

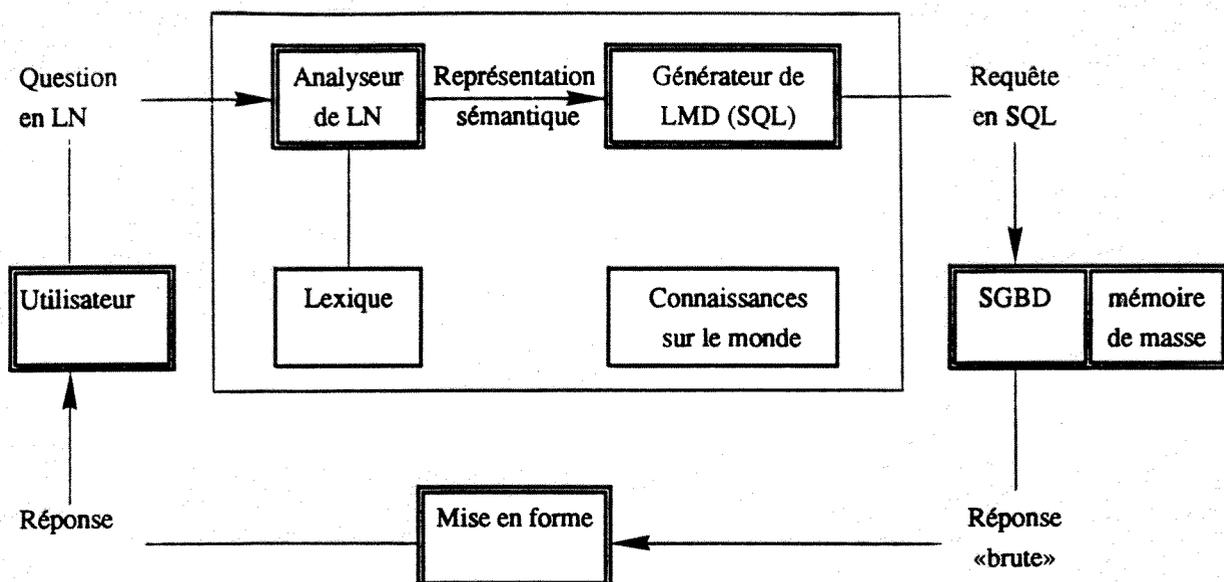
- . ils sont limités à un domaine bien précis et souvent à des requêtes triviales ;
- . la compréhension des requêtes et l'analyse automatique du langage naturel sont compliquées, ce qui oblige les concepteurs à prévoir dans la majorité des cas, un dialogue pour lever les ambiguïtés, faire des mises au point, compléter les requêtes, etc...

II . 2 . 1 . 3 Modèle d'architecture

Le concept d'interface inclu :

- l'identification des éléments dont elle fait la conjonction, le SGBD d'une part et l'usager de l'autre ;
- un langage qui permet l'échange de message entre deux éléments ;
- des règles syntaxiques qui définissent quelles sont les phrases correctes, exprimées avec le langage de l'interface ;
- des règles de dialogue qui définissent les enchaînements corrects entre les phrases du langage de l'interface ;
- un univers de discours commun aux deux éléments qui assure une interprétation non ambiguë des phrases échangées.

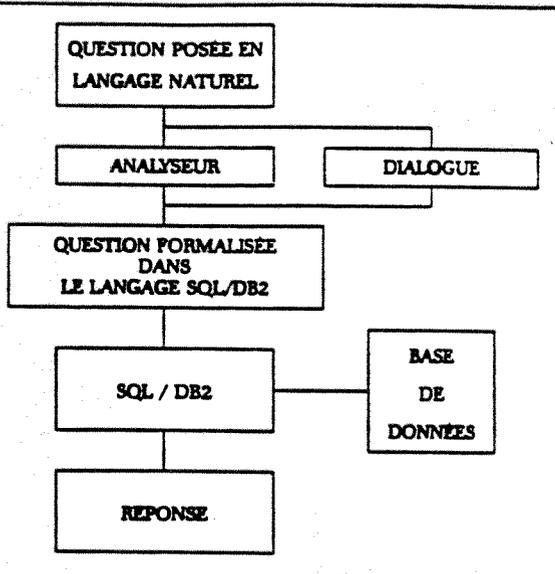
Présentation de la couche dite *noble* de l'interface :



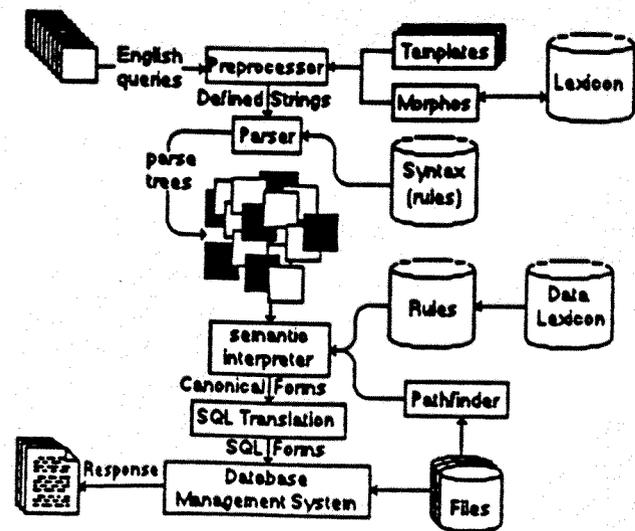
Ce modèle d'architecture est celui utilisé dans les interfaces :

- SYSTEM X: ref [18-CERCONE-1993] / [72-McFETRIDGE-1988]
- NAT: ref [65-COCH-DI-YACOVO-1991]
- SAPHIR: ref [36-NORMIER-1982] / [42-SAPHIR-1987]

II . 2 . 1 . 4 Exemples d'architectures⁴



architecture de SAPHIR



architecture de System X

II . 2 . 1 . 5 Les références obtenues

- SIMPLA: ref [12-ARITA-1994]
- TEXT-IS: ref [27-KNOWLES-1994]
- IDQS: ref [44-SHOW-1993]
- NLQI: ref [64-CHAU-1993]
- PC/ICON: ref [14-BALLOU-1993]
- TELI: ref [57-BALLARD-1993]
- ATMI: ref [62-BURTON-1992]
- TLG: ref [59-BRYANT-1992]
- SESAME: ref [80,81-SABBAGH-1990]
- NLI: ref [60-BURTON-1989]
- NQL approach: ref [75-OBERMEIER-1988]
- RM/T: ref [25-KAO-1988]
- MENTAT: ref [76-OGONOWSKI-1987]

⁴ Architectures présentées dans la thèse ref [07-PUJO-1989].

II . 2 . 2 Présentation sommaire d'autres types d'interfaces

II . 2 . 2 . 1 Les interfaces graphiques

II . 2 . 2 . 1 . 1 Introduction

La requête est construite spécialement, avec une utilisation réduite du clavier mais par la souris avec navigation par îcones. L'ouvrage ref [02-MEINADIER-1991] a été proposé pour permettre de faire un parallèle entre les interfaces graphiques orientées utilisateurs et celles concernant l'interrogation des bases de données relationnelles. Ces interfaces sont aussi appelées à *syntaxe bi-dimensionnelles*.

II . 2 . 2 . 1 . 2 Les références obtenues

GUI: ref [64-CHAU-1993]
GRAOULA: ref [46-SOKUT-1993]
ACORD: ref [01-BES-1992]
Graphical User Interface ..: ref [55-ZINN-1992]
RUITALK: ref [09-SAKER-1992]

II . 2 . 2 . 2 Les interfaces relations universelles

II . 2 . 2 . 2 . 1 Introduction

Une tendance actuelle vise à bâtir des interfaces pour non-informaticiens permettant de masquer la vision multi-relationnelle de la base de données, et de faciliter la formulation de requêtes dans les langages relationnels. Cette tendance utilise une approche fondée sur le concept "relation universelle". L'approche relation universelle ne constitue pas une théorie stable mais plutôt une hypothèse expérimentale, cette hypothèse utilise le principe commun pour la formulation des requêtes de type :

```
SELECT < attribut résultat >  
FROM < table >  
WHERE < condition >
```

II . 2 . 2 . 2 . 2 Les références obtenues

UQL: ref [68-HYOTYNIEMI-1992]
SYSTEM U: ref [51-VARDI-1988]
FIDL: ref [51-VARDI-1987]

II . 3 La génération de requêtes à partir du langage naturel

II . 3 . 1 Introduction de la notion de paraphrase

Dans cette partie nous présentons un élément de départ pour la génération de requêtes qu'est le paraphrasage ainsi que des produits de générations de requêtes.

La paraphrase a pour but essentiel de valider le processus de communication dans le sens homme-machine. Ce processus, en effet, peut échouer pour deux raisons principales :

- la requête est incorrectement formulée ;
- la requête correcte, est mal interprétée par l'interface.

Le deuxième cas est, en général plus fréquent que le premier ; l'utilisateur a une meilleure performance linguiste que son interlocuteur électronique.

La paraphrase peut révéler à l'utilisateur certaines erreurs⁵ que contient sa question. Dans les deux cas, paraphraser la question initiale est nécessaire. L'obtention du résultat ne suffit pas, sauf pour une requête très simple, s'assurer que la question a été bien comprise. Il s'avère frustrant, dès que la question est complexe, de n'obtenir aucune information sur le traitement effectué par la machine. Le crédit que l'utilisateur peut accorder aux informations fournies est en rapport avec la qualité des justifications apportées par le système. Ce phénomène est d'autant plus marqué que la réponse revêt un caractère important aux yeux du questionneur.

Pour résumer, le paraphrasage permet de passer d'une question en langage naturel vers un langage pseudo-naturel.

II . 3 . 2 Exemple de la paraphrase sur SAPHIR

Les concepteurs de SAPHIR rapportent que les utilisateurs se sentent peu responsables de la qualité des réponses qu'ils obtiennent. Si l'information délivrée est satisfaisante, le travail du système est incriminé alors que, parfois, le problème résulte d'une mauvaise formulation de la requête initiale. Quand les utilisateurs maîtrisent parfaitement la nature ou la structure des données, la paraphrase fournie ne permet pas toujours de révéler ce défaut ; seul le détail des opérations effectuées sur la base peut mettre en évidence ce problème. Pour tenter de résoudre ce problème, SAPHIR propose à l'utilisateur d'afficher entre la paraphrase, le programme SQL qu'il a généré.

Exemple de paraphrase de SAPHIR, page suivante :

⁵Dans cette analyse, on ne détaillera pas le traitement des erreurs de requêtes, mais on pourra se reporter à la ref [07-PUJO-1989] qui traite de ce sujet.

SAPHIR : POSEZ VOTRE QUESTION

Profs de sexe féminin qui assurent des cours dans un établissement de la ville de Colmar ?

```
.....Structured Query Language.....  
SELECT B.NOM,-  
        F.MATIERE,-  
        I.NOM-  
FROM PROF B,-  
        COURS F,-  
        ETABLI I,-  
        CLASSE O-  
WHERE B.SEXE = 2-  
        AND B.NOPR = F.NOPR-  
        AND O.NOCL = F.NOCL-  
        AND O.NOET = I.NOET-  
        AND I.VILLE = COLMAR
```

On peut ainsi se rendre compte du bien-fondé de l'évaluation de leur question et détecter, le cas échéant, les erreurs commises dans l'expression de leur recherche d'informations. Hélas, l'affichage du programme SQL n'est bien évidemment d'aucun secours pour les utilisateurs non-informaticiens.

II . 3 . 3 Les références obtenues

DATENBANK-DIALOG: ref [83-TROST-1992]
HSQL: ref [71-LUJNGBERG-1991]
INRS-ATIS: ref [63-CHANGWEN-1992]
Interface de correction de requêtes: ref [07-PUJO-1989]
KDA: ref [84-WU-1989]
LINUS: ref [66-ELLENA-1989]
RELATIONAL QUERY: ref [56-AMANO-1993]
REMIT: ref [31-LOWDEN-1987]

III . CONCLUSION

Dans notre étude, on remarque la prédominance des interfaces en langage naturel par rapport aux autres types présentés. Ceci s'explique aisément par leur rôle de pionnier dans ce domaine. L'avènement de l'intelligence artificielle, notamment les systèmes experts ont été un moteur qui a favorisé le développement de ces interfaces. Par définition, le langage naturel doit augmenter sensiblement la convivialité mais quant à l'efficacité de la recherche il s'agit là, d'un autre point bien distinct sur lequel le langage naturel doit encore faire ses preuves.

Les interfaces en langage naturel ont évolué aujourd'hui vers une approche multimédia intégrant :

- des méthodes de reconnaissance de la voix ;
- des outils d'interprétation des expressions ;
- des techniques de traitement d'image.

Lorsque la nature suffisamment simple de l'information et la faible complexité des questions le permettaient, les interfaces graphiques étaient les plus adaptées, celles-ci ont commencé à être commercialisées bien après les interfaces en langage naturel, il s'avère que ces produits n'étaient pas considérés comme des outils complémentaires mais plutôt comme des rivaux aux interfaces "classiques".

Pour ce qui est des interfaces relations universelles, elles n'ont pas eu autant de succès ; peut-être parce que leur concept de base ne se fonde pas sur une théorie stable, mais plutôt une hypothèse expérimentale.

Enfin, rappelons qu'un nombre important d'outils gravitant autour de ces interfaces ont vu le jour :

- les analyseurs de forme logique pour le développement de langage pseudo-naturel ;
- les systèmes de mise à jour automatique de la base de connaissance (bases "dynamiques") ;
- les bases de données expertes pour la manipulation d'objets complexes ;
- le développement en mode conversationnel de requêtes ;
- la portabilité des interfaces sur différentes plate-formes ;
- le développement de logiciels d'optimisation et de correction de requêtes ;
- les générateurs d'interfaces...

BIBLIOGRAPHIE

I FORMAT DES NOTICES

Normalisation :

Les références bibliographiques ont été présentées selon la dernière norme **AFNOR Z 44 - 005** de Décembre 1987.

ISSN :

La majorité des articles de périodiques recensés comportent leurs numéros d'ISSN, sauf pour six revues dont cinq américaines (Technological, Acm, Digital review, Info world, PC week) et une française (Logiciels et services).

ISBN :

Aucune difficulté ne s'est posée pour retrouver les numéros d'ISBN des deux ouvrages cités.

Pagination :

La plupart des références obtenues à partir de la base UNCOVER n'indiquaient pas le détail de la pagination, nous nous sommes limités à présenter le numéro de page de début d'article.

II CLASSEMENT DES REFERENCES

La diversité des documents obtenus, nous a amené à un classement des quatre vingt quatre références selon deux critères de façon à fournir une bibliographie la plus lisible possible.

Ceux sont en ordre de priorité :

- par type de documents : *monographies, rapports de recherches, thèses, articles de périodiques et actes de congrès* ;
- par ordre alphabétique d'auteurs.

La clé de recherche par auteur a été préférée à un classement chronologique pour deux raisons essentielles :

- la consultation de la bibliographie est plus aisée sur une recherche par nom d'auteurs que sur la date de publication d'un article ;
- un classement par date, notamment lors de la recherche sur les dernières publications, risque de défavoriser la consultation d'articles parus dans la période de 1988 à 1990, où de nombreux articles et manifestations ont traité de notre sujet.

Ci-dessous, un tableau comparatif révélateur :

Années	1987	1988	1989	1990	1991	1992	1993	1994
Nb.d'articles	5	11	14	14	8	11	12	7

III PRESENTATION DE LA BIBLIOGRAPHIE

III . 1 Monographies

01

BES, G.G. ; GUILLOTIN, T.- *A natural language and graphics interface : results and perspectives from the ACORD project*, Berlin, Heidelberg, New-york : ed. Springer Verlag, 1992, 176 p.
ISBN 3-540-55675-3

02

MEINADIER, J.P.- *L'interface utilisateur pour une informatique plus conviviale*. Informatique et Stratégie, Paris : ed. Dunod, 1991, 222 p.
ISBN 2-10-000160-4

III . 2 Rapports de recherches

03

BUCHBERGER, E ; TROST, H.- *DIALOG-databank : communication with databanks in the German language*. Report Number : TR-89-7, 1989.

04

PICHAT, E. ; SAKER, D.- *Interrogation graphique des bases de données*. Rapport de recherche LIGIA, Université Claude Bernard Lyon I, 1993, 19 p.

III . 3 Thèses

05

ELLUL-COMMARE, A.- *Interface langage naturel / base de données dans le domaine des petites annonces informatiques*. Thèse de doctorat, ENST de Paris, 1987, 162 p.

06

LAKHAL, L.- *Contribution à l'étude des interfaces pour non-informaticiens dans la manipulation de bases de données relationnelles*. Thèse de doctorat d'état, Université Sciences de Nice, 1986, 117 p.

07

PUJO, P.- *Développement d'une interface conviviale, pour l'interrogation en langage naturel d'une base de données, avec utilisation des concepts et des moyens de l'intelligence artificielle*. Thèse de doctorat, Université Paris XI, 1989, 273 p.

08

SABATIER, P.- *Contribution au développement d'interfaces en langage naturel*. Thèse de doctorat d'état, Université Paris VII, 1987, 193 p.

09

SAKER, D.- *L'interrogation des bases de données relationnelles assistée par le graphe sémantique normalisé.* Thèse de doctorat, Université Claude Bernard Lyon I, 1993, 108 p.

III . 4 Articles de périodiques

10

ALLUISI, J.C.; STAGGS, J.L.- *A Student Information System Based on Relational Technology.* Technological horizons in education , 1990, vol. 18, n° 4, p. 72.

11

ANICK, P.G.- *Integrating natural language processing and information retrieval in a troubleshooting help desk.* IEEE Expert, 1993, vol. 8, n° 6, p. 9-18.
ISSN 0885-9000

12

ARITA, S. ; NISHIMURA, K. ; SHIMAZU, H.- *Multimodally natural language interface SIMPLA.* NEC Research and Development, 1994, vol. 35, p. 421-425.
ISSN 0547-051X

13

ASBRAND, D.- *Realty firm takes do-it-yours et SQLef approach to database.* InfoWorld, 1994, vol. 16, n° 18, p. 85-86.
ISSN 0199-6649

14

BALLOU, M.C.- *PC/Icon : English-speaking tool.* Computerworld, 1993, vol. 27, n° 45, p. 76-77.
ISSN 0010-4841

15

BIRKHEAD, E.- *The natural query program.* DEC Professional, 1989, vol. 8, n° 9 , p. 74-78.
ISSN 0744-9216

16

BLEED, R.- *Natural Language System Tracks High School Students At Risk.* Technological horizons in education, 1990, vol. 18, n° 2, p. 77.

17

BOZMAN, J.S.- *Oracle tous Easytalk language interface for VAXs.* Computerworld, 1989, vol. 23, n° 34, p 34-35.
ISSN 0010-4841

18

CERCONE, N. ; JIAWEI, HAN ; McFETRIDGE, P. ; ET AL.- *SystemX & DBLEARN: easily getting more from your relational database.* Integrated Computer-Aided Engineering, 1993, vol. 1, n° 1, p. 311-339.
ISSN 1069-2509

19

CLIFFORD, J. ; CROCKER, A. ; TUZHILIN, A.- *On Completeness of Historical Relational Query languages.* *Acm transactions on databases systems*, 1994 ,vol. 19, n° 1, p. 64.
ISSN 0362-5915

20

CZEJDO, B. ; ELMASRI, R. ; RUSINKIEWICS, M.- *A Graphical Data Manipulation Language for an Extended Entity-Relationship Model.* *Computer*, 1990, vol. 23, n° 3, p. 26.
ISSN 0738-9213

21

DECLÉ, J.P.- *Les interfaces homme-machine en langage naturel dans l'interrogation des bases de données relationnelles.* *Information, documentation, transfert des connaissances*, 1994.
ISSN 1140-6992

22

DOLER, K.- *AI techniques are behind the approaches being used to deal with SQL's syntax.* *PC Week*, 1988, vol. 5, n° 17, p. 17-19.

23

FREYTAG, J.C. ; GOODMAN, N.- *On the Translation of Relational Queries into Iterative Programs.* *Acm transactions on database systems*, 1989, vol. 14, n° 1, p. 1.

24

FROST, R.A ; SABA, W.S.- *A database interface based on Montague's approach to the interpretation of natural language.* *International journal of man machine studies*, 1990, vol. 3, n° 2, p. 149-176.
ISSN 0020-7373

25

KAO, M. ; CERCONE, N. ; LUK, W.S.- *Providing quality responses with natural language interfaces : the null value problem.* *IEEE Transactions on Software Engineering*, 1988, vol. 14, n° 7, p. 959-984.
ISSN 0098-5589

26

KINUKAWA, H.- *A natural language interface processor based on the hierarchical-tree structure model of relation tables.* *Journal of Information Processing*, 1988, vol. 11, n° 2, p. 83-91.
ISSN 0387-6101

27

KNOWLES, A.- *Enterprise database matures : Tex-is server package does real-time search via English queries.* *PC Week*, 1994, vol. 11, n° 15, p. 66.
ISSN 0740-1604

28

KUZNETSOV, S.D.- *Logical Query Optimization in Relational Database Management Systems*. Programming and computer software, 1989, vol. 15, n° 6, p. 271-272.
ISSN 0361-7688

29

KUZNETSOV, S.D.- *Generation of Optimal Query Execution Schemes in Relational Database Management Systems*. Programming and computer software, 1990, vol. 16, n° 2, p. 47.
ISSN 0361-7688

30

LATAMORE, G.B.- *Getting into the mainframe*. Personal Computing, 1988, vol. 12, n° 4, p. 81-83.
ISSN 0192-5490

31

LOWDEN, B.G.T. ; DE ROECK, A.N. ; PHILIHIPPS, D.J. ; ET AL.- *General purpose natural language interface design and application as a database front end*. ICL Technical Journal International Computers Limited, 1987, vol. 5, n° 4, p. 763-777.
ISSN 0142-1557

32

MASUD, S.A.- *DBMS helps NASA keep on top of task orders*. Government Computer News, 1990, vol. 9, n° 21, p. 37-38.
ISSN 0738-4300

33

MC LEAN, S. ; WEISE, C.- *DIGRESS : a deductive interface to a relational database*. Journal of the american society for information science, 1991, vol. 42, n° 1, p. 49-57.
ISSN 0002-8231

34

MELNITSKY, S.- *SQL as a second language*. Digital Review, 1990, vol. 7, n° 13, p. 27-31.
ISSN 0739-4314

35

NORMIER, B.- *Interfaces en langage naturel avec des bases de données*. Télécom, 1991, n° 87, p. 36-39.
ISSN 0040-2478

36

NORMIER, B. ; BOURGUINON, J.M. ; BUZENAT, B.- *SAPHIR : système d'interrogation des bases de données relationnelles*. TA Informations, 1982, n° 2, p. 32-58.
ISSN 0039-8217

37

OTT, N.- *Aspects of the automatic generation of SQL statements in a natural language query interface*. Information Systems, 1992, vol. 17, n° 2, p. 147-159.
ISSN 0306-4379

38

OZSU, M. ; TAMER, D.J.- *Finding heuristics for processing selection queries in relational database systems.* Information systems, 1990, vol. 15, n° 3, p. 359.
ISSN 0248-8590

39

PETROVIC, S.- *Providing help in a natural language query interface to relational databases.* Yugoslav journal of operations research, 1992, p. 207-218.
ISSN 7127-3026

40

PUJO, P. ; LANGELET, G.- *Aide intelligente pour la correction des requêtes en langage naturel à une base de données.* Technique et sciences informatiques, 1990, vol. 9, n° 3, p. 217-230.
ISSN 0752-4072

41

PUTTRE, M.- *EasyTalk makes access as easy as abc : The Hartford pioneers natural language programming that bypasses experts, helps novices.* Information Week, 1989, n° 215, p. 25-27.
ISSN 8750-6874

42

SAPHIR.- *Le langage pour DB2 et SQL.* Logiciels et services, 1987, n° 64, p. 37-39.

43

SCEMAMA, C.- *Langage naturel et consultation des banques de données.* Bases, 1989, n° 36, p. 5-6.
ISSN 0765-1325

44

SHOW, G.Y. ; HINNY-KONG ; LIN, W.K.- *Intelligent user interface to SQL-based database system.* Engineering Applications of Artificial Intelligence, 1993, vol. 6, n° 4, p. 307-316.
ISSN 0952-1976

45

SKOTNICZNY, Z.- *Query by Forms: user-oriented relational database retrieving system and its application in analysis of experiment data.* Computer physics communications, 1989, vol. 57, n° 1-3, p. 225.
ISSN 0010-4655

46

SOCKUT, G.H. ; BURNS, M.S. ; MALHOTRA, A. ; ET AL.- *GRAOULA : a graphical query language for entity-relational databases.* Data and knowledge engineering, 1993, vol. 3, n° 2, p. 171-202.
ISSN 0169-9023

47

STREHLO, K.- *Database front end for Windows speaks users' native language.* InfoWorld, 1993, vol. 15, n° 6, p. 1.

48

TAKAHASHI, Y.- *Fuzzy Database Query Languages and Their Relational Completeness Theorem.* IEEE transactions on knowledge and data engineering, 1993, vol. 5, n° 1, p. 122.
ISSN 1041-4347

49

TSENG, F.S.C. ; CHEN, A.L.P. ; WEI-PANG ,Y.- *On mapping natural language constructs into relational.* Data & Knowledge Engineering, 1992, vol. 9, n° 1, p. 97-118.
ISSN 0169-023X

50

UKOVSKY, S.- *Users need just plain English for Britton Lee DBMS queries.* Digital Review, 1989, vol. 6, n° 9, p. 25-26.

51

VARDI, M.Y.- *The universal-relation data model for logical independence.* IEEE Software, 1988 , vol. 5, n° 2, p. 80-86.
ISSN 0740-7459

52

VARINARD, C.- *Interface en langue naturelle pour l'interrogation d'une base de donnees.* Les Cahiers Lyonnais de Recherche en Gestion, 1990, vol. 11, p. 26.
ISSN 0606-2428

53

WANG, F.- *Towards a natural language user interface : an approach of fuzzy query.* International journal of geographical information, 1994, vol. 8, n° 2, p. 143.
ISSN 0269-3798

54

WHANG, K.Y. ; KRISHNAUMURTHY, R.- *Query Optimization in a Memory Resident Domain Relational Calculus Database System.* Acm transactions on database systems, 1990, vol. 15, n° 1, p. 67.

55

ZINN, T.K.- *Enabling technologies build bridge to 21st century.* Computers in Healthcare, 1992, vol. 13, n° 12, p. 28-31.
ISSN 0745-1075

III . 5 Actes de congrès

56

AMANO, H. ; KAMBAYASHI, Y.- *Relational query formulation by pseudonatural language text manipulation.* In Proceedings of the Third International Symposium on Database Systems for Advanced Applications, Taejon, South Korea, Apr. 1993.- Singapore: World Scientific, 1993, p. 11-18.

57

BALLARD, B.- *TELI : a powerful, easily adaptable natural language question answering system for information retrieval.* In Online '87 Conference Proceedings, Anaheim, USA, Oct. 87.- Weston, USA : Online, 1993, p. 9-13.

58

BRENNAN, S.E.- *The multimedia articulation of answers in a natural language database query system.* In Second Conference on Applied Natural Language Processing, Austin, USA, Feb. 1988.- Morristown, USA : Assoc. Comput. Linguistics, 1988, p. 1-8.

59

BRYANT, B.R. ; PAN, A.- *Two-level grammar: a functional/logic query language for database and knowledge-base systems.* In International Conference Logic Programming and Automated Reasoning LPAR '92, St.Petersburg, Russia, July 1992.- Berlin, Germany : Springer-Verlag, 1992, p. 78-83.

60

BURTON, A. ; STEWARD, A. ; DAVENPORT, C.- *An adaptive natural language interface to management information.* In Proceedings of the 11th BCS IRSG Research Colloquium on Information Retrieval, Huddersfield, UK, July 1989.- Huddersfield : Huddersfield Polytech., 1989, p. 26-34.

61

BURTON, A. ; STEWARD, A.- *The design and implementation of a context sensitive natural language interface to management information.* In Proceedings of the IFIP TC 13 Third International Conference on Human Computer Interaction INTERACT'90, Cambridge, UK, Aug. 1990.- Amsterdam, Netherlands : North-Holland Publ. Co., 1990, p. 795-800.

62

BURTON, A. ; STEWARD, A.P.- *A natural language interface to management information.* In International Conference Practical Applications of Prolog, 1992, London, UK, Apr. 1992.- Dorking, UK : Practical Applications of Prolog, 1992, vol. 1, p. 1-23.

63

CHANGWEN, Y. ; O'SHAUGHNESSY, D.- *The INRS ATIS system and its pragmatic analysis.* In Proceedings of the First Singapore International on Intelligent Systems, SPICIS '92-, Singapore, Sept.-Oct. 1992.- Singapore : Singapore AI Centre, 1992, p. 353-358.

64

CHAU, R.- *Natural language interfaces for integrated network management.* In Proceedings of 5th International Conference on Human-Computer Interaction, Orlando, USA, Aug. 1993.- Amsterdam, Netherlands : Elsevier, 1993, vol. 2, p. 368-372.

65

COCH-DI-YACOVO, J. ; CRISPINO, G. ; CUKIERMAN, D. ; ET AL.- *Nat-multiling tools for multilingual interfaces with data bases.* In RIAO 91 : recherche d'information assistée par ordinateur, Barcelona, Spain, Apr. 1991.- Marseille : Centre des hautes études internationales d'informatique documentaire, 1991, vol. 1, p. 514-525.

66

ELLENA, C. ; ERRICO, M. ; MAROTTA, G. ; ET AL.- *LINUS : an Italian interface for relational databases.* In Annual Conference Proceedings, Trieste, Italy, Oct. 1989.- Milan, Italy : A.I.C.A., 1989, vol. 1, p. 525-539.

67

ELSEVIER- *Intelligent Information Access.* In Proceedings of the BANKAI Workshop, Brussels, Oct. 1991. Amsterdam, Netherlands : Elsevier, 1992, p. 187.

68

HYOTYNIEMI, H. ; LRHTOLA, A.- *A universal relation database interface for knowledge based.* In Proceedings of the Second International Symposium, Database Systems for Advanced Applications'91, Tokyo, Japan, Apr. 1991.- Singapore : World Scientific, 1992, p. 84-88.

69

JONES, K.P.- *Prospects for Intelligent Retrieval.* In Proceedings of a Conference Informatics 10, Cambridge, UK, Mar. 1989.- London, UK : Aslib, 1990.

70

JOSEPH, S.W. ; ALELIUNAS, R.- *A knowledge-based subsystem for a natural language interface to a database that predicts and explains query.* In Proceedings Seventh International Conference on Data Engineering, Kobe, Japan, Apr. 1991.- Los Alamitos, USA : I.E.E.E. Comput. Soc. Press, 1991, p. 80-87.

71

LJUNGBERG, J.- *Paraphrasing SQL to natural language.* In Proceedings of a Conference. RIAO '91, Intelligent Text and Image Handling, Barcelona, Spain, Apr. 1991. - Amsterdam, Netherlands : Elsevier, 1991, p. 790-808.

72

McFETRIDGE, P. ; HALL, G. ; CERCONE, N. ; ET AL.- *System X: a portable natural language interface.* In Proceedings of the Seventh Biennial Conference of the Canadian Society for Computational Studies of Intelligence, Edmonton, Canada, June 1988.- Toronto, Canada : University Alberta Press, 1988, p. 30-38.

73

McFETRIDGE, P. ; HALL, G. ; CERCONE, N. ; ET AL.- *Knowledge acquisition in System X: natural language interface to relational databases.* In Proceedings International Computer Science Conference '88. Artificial Intelligence : Theory and Applications, Hong Kong, Dec. 1988.- Hong Kong : I.E.E.E. Computer. Soc., 1988, p. 604-610.

74

MEHL, S.- *A plea for better crossfile searching facilities among databases and databases hosts.* In International onliner information meeting, London, UK, Dec. 1988.- Oxford, UK : Learned Information Europe, 1988, vol. 1, p. 35-40.

75

OBERMEIER, K.K.- *The NLQ approach to database retrieval.* In Proceedings of the IEEE, 1989, National Aerospace and Electronics Conference NAECON, Dayton, USA, May 1988.- New-York, USA : I.E.E.E., 1989, vol. 3, p. 1370-1373.

76

OGONOWSKI, A.- *MENTAT : an intelligent and cooperative natural language DB interface in Avignon 87*. In Journées Internationales. Les Systèmes Experts et leurs Applications, Avignon, Mai 1987.- Paris : Ass. Française pour la Cybernétique Economique et Technique, 1987, p. 1271-1283.

77

PEQUEGNAT, C. ; AGUIRRE, J.L. ; BATTANDIER, J.L. ; ET AL.- *Interface en langue naturelle pour l'interrogation d'une base de données relationnelles in Avignon 88*. In Journées Internationales. Les Systèmes Experts et leurs Applications, 1988.

78

RAMOS, H.- *IQL, Interface graphique pour base de données relationnelles*. In Actes du séminaire sur les BD 91, Alger, Algérie, 1991, p. 365-385.

79

RAYNER, M. ; ALSHAWI, H.- *Deriving database queries from logical forms by abductive definition expansion*. In Proceedings of 3rd Conference on Applied Natural Language Processing, Trento, Italy, Mar.-Apr. 1992.- Morristown, USA : Ass. Comput. Linguistics, 1992, p. 1-8.

80

SABBAGH, S.- *SESAME : an application of entity relationship models to a natural language user interface - Entity-Relationship Approach : the Core of Conceptual Modelling*. In Proceedings of the Ninth International Conference, Lausanne, Switzerland, Oct. 1990.- Amsterdam, Netherlands : North Holland Pub. Co., 1991, p. 319-331.

81

SABBAGH, S.- *SESAME : an interface generator for relational databases*. In Tenth International Workshop. Expert Systems and their Applications. Specialized Conference : Natural Language Processing and its Applications, Avignon, May-June 1990.- Nanterre : EC2, 1990, p. 35-50.

82

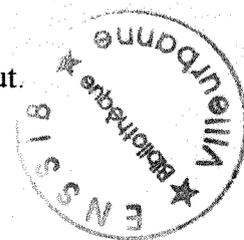
SANAMRAD, M.A. ; BRETAN, I.- *IBM SAA Language Access : a large-scale commercial product implemented in PROLOG*. In Practical Applications of Prolog. International Conference, 1992, London, UK, Apr. 1992.- Dorking, UK : Practical Applications of Prolog, 1992, vol. 1, p. 3-14.

83

TROST, H. ; HEINZ, W. ; MATIASEK, J. ; ET AL.- *Datenbank-DIALOG and the relevance of habitability*. In Proceedings of 3rd Conference on Applied Natural Language Processing, 1992, Trento, Italy, Mar.-Apr. 1992.- Morristown, USA : Ass. Computer Linguistics, 1992, p. 241-242.

84

WU, X. ; TANAKA, M. ; ICHIKAWA, T.- *KDA : a knowledge-based database assistant*. In Proceedings Fifth International Conference on Data Engineering, Los Angeles, USA, Feb. 1989.- Washington, USA : I.E.E.E. Comput. Soc. Press, 1989, p. 402-409.



BIBLIOTHEQUE DE L'ENSSIB



9666732