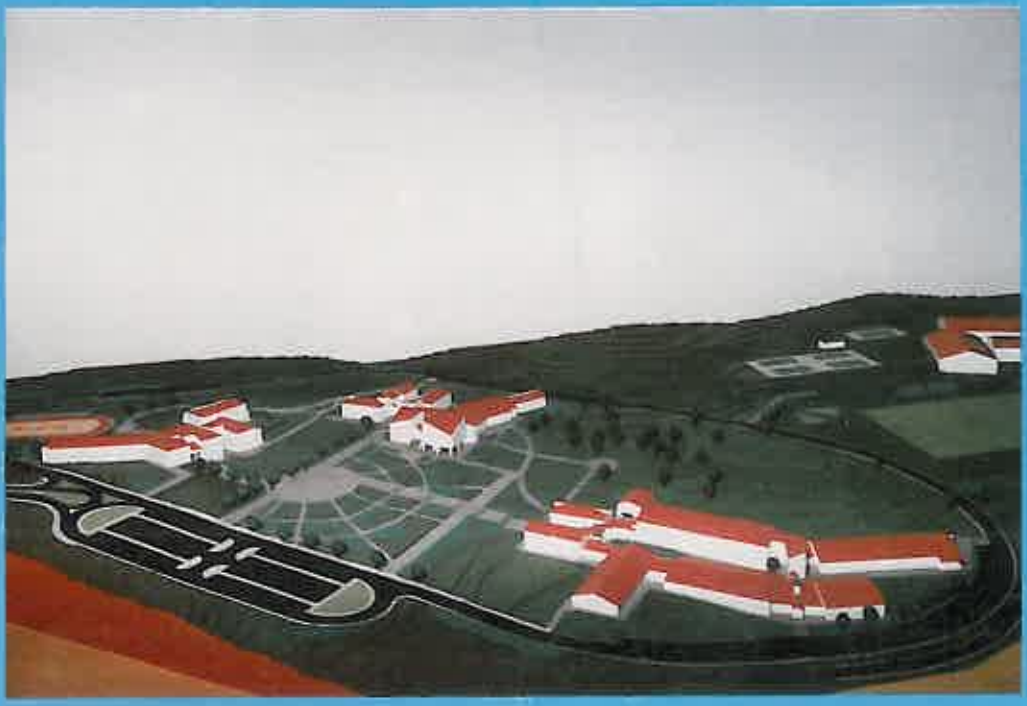
EDUCAÇÃO e TECNOLOGIA



Revista do Instituto Politécnico da Guarda

"EDUCAÇÃO E TECNOLOGIA"
Revista do Instituto Politecnico da Guarda

DIRECTOR: João Bento Raimundo

REDACÇÃO: Rua Comandante Salvador do Nascimento
Telef. 21634 6300 GUARDA

PROPRIEDADE: Instituto Politécnico da Guarda

EXECUÇÃO GRÁFICA: Secção de Reprografia do IPG

Depósito Legal N.º 17.891/87

Nº 5 / Setembro / 89

ABERTURA PARA O MUNDO ...

"Português que viva apenas para Portugal, como acho queria o Velho do Restelo, não tem significado algum nem vale a pena existir no mundo; temos de viver para o universo, ou seremos inúteis".

Agostinho da Silva

Sempre defendemos a formação integral do indivíduo. Tal significa, para nós, em termos globais, o crescimento perante conhecimentos gerais e específicos; o acordar das potencialidades de cada um; a afirmação do indivíduo perante ele próprio, em primeiro lugar, perante os outros e o mundo, depois; o, já tantas vezes referido, saber, saber fazer, saber ser; enfim, um caminhar efectivo para a realização e para a felicidade.

O presente número, o quinto, de "Educação e Tecnologia", enquanto "um espaço aberto", objectivo — génese da sua existência e da sua afirmação — na linha do que atrás referimos, inclui já a participação de professores de Instituições ligadas ao Instituto Politécnico da Guarda pelo Programa Erasmus. Isto constitui um sinal evidente da cooperação que, a vários níveis, há alguns meses atrás, foi acordada em protocolos com Bayonne, Brighton, Coventry, Créteil, Pau e Salamanca.

Este aprofundamento de relações entre instituições europeias de ensino superior veio favorecer a vivência do espírito comunitário e imprimir nos alunos a consciencialização do conceito da nova Europa da cultura e dos cidadãos.

Defendemos e prosseguimos um caminho de abertura para o mundo das coisas, das pessoas e do saber, numa perspectiva integradora em que a verdadeira dimensão do humano se procure, se veja e se consubstancie na efectiva comunhão do universal.

João Bento Raimundo

Presidente da C. I. do
Instituto Politécnico da Guarda

A MULTI-MEDIA INFORMATION BASE FOR THE INFORMATION AND CAREER GUIDANCE OF SECONDARY SCHOOL PUPILS

- THE AQUEDUCT SYSTEM -

Jean-Paul Anton, Françoise Labrieu*, Pantxika Dagorret **

ABSTRACT:

AQUEDUCT is a general public multimedia system, mainly designed for secondary school pupils. It provides them with detailed and coherent informat on about jobs and teaching.

The system currently includes three components:

- a "text" component which associates an adaptable information structure, automatic indexing procedures and natural language querying facilities;
- an "image" component, handled by videodisk, used both as an illustration of the text and as a help to it's querying;
- a "graphic" component, allowing to display the set of paths leading from a basic level to a higher level job or training.

This project was developed in collaboration with the Regional Bureau of the "Office National de l'Information Sur les Enseignements et les Professions" (DRONISEO) of BORDEAUX, the "Office Régional de l'Education Permanente" (OREP) of PAU, and with the participation of the Aquitaine Information and Career guidance Centers.

* Université Paul Sabatier.
** Institut Universitaire de Technologie de Bayonne.

I - INTRODUCTION

The mission of C.I.O.s (Centre d'Information et d'Orientation = Information and Career guidance Center), located in most medium-size cities, consists in:

- collecting and distributing all information ralated to the teachings provided by secondary and university level education;
- supplying schoolchildren, secondary school pupils and students advice for choosing such or such type of studies or jobs;
- developing relations with the business world and employment environment;
- facilitating professional integration.

These actions cannot be achieved without the cooperation of various partners (National Education, employment agencies, and institutions for professional and adult continuing education, under the supervision and coordenation of the ONISEP (Office National d'Information Sur les Enseignements et les Professions) and more particulary of each local bureau wich ensures the production and updating of local school and professional information in its own area.

The set of very heterogeneous information they have at their disposal (texts, films, photographs, graphics, ...) must be easily consulted and understood by a large number of different individuals (educators, secondary school pupils, parents, and so on...).

Therefore, the local bureau of the ONISEP in Bordeaux has chosen to implement a Multimedia Information Base about teaching and professions in Aquitaine: the AQUEDUCT project (AQUitaine EDUCaTion).

AQUEDUCT was developed to meet the following requirements:

- capacity to manage a huge bulk of non structured information: texts, graphics, fixed and motion pictures;
- flexibility to facilitate:
  . consultations by any occasional user whether a computer specialist or not;
  . the management and updating of the base by C.I.O. members.
- decent response times;
- compatibility of supports.

The information will be obtained from "multimedia stations" that may be autonomous or controlled by a telematica server, located in C.I.O. offices or educational institutions.

The following pages introduce the general features of the system and its main functions. Some more theoretical and technical aspects are to be found in: [ANT 85] , [ANT 87], [CHR 85-a], [CHR 85-b], [DAG 87].

II - BASE CONTENTS

The base documents (or cards) include texts, graphics and fixed or motion pictures. Texts and graphics define the main characteristics (type of work, working conditions, ...) as well as the training required for job described on the picture.

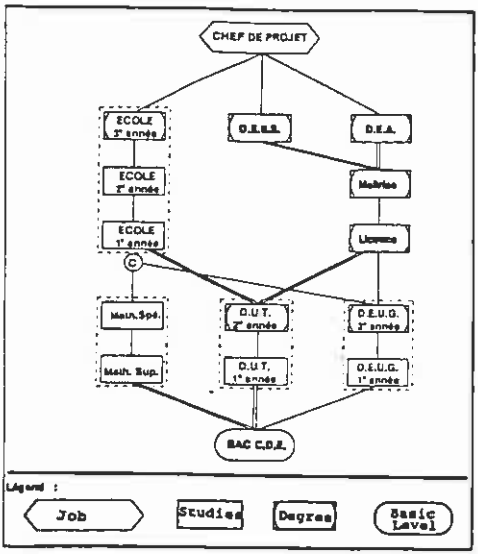


Fig 1: Examples of base texts and graphics.

III - BASE ORGANIZATION

The base is made up of a set of cards (or documents). Each one contains a title which sumarizes the main features of a job, and pages that provide further information about it.

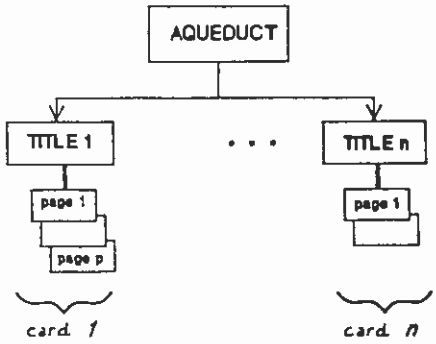


Fig. 2: Logical diagram of the base.

This structure is completed with references, which materialize the logical links between the documents. Thus making consultation quicker and more flexible: the user may obtain additional information to the contents of a card by directly displaying the documents referenced on it. [CHR 85] [VEL 84].

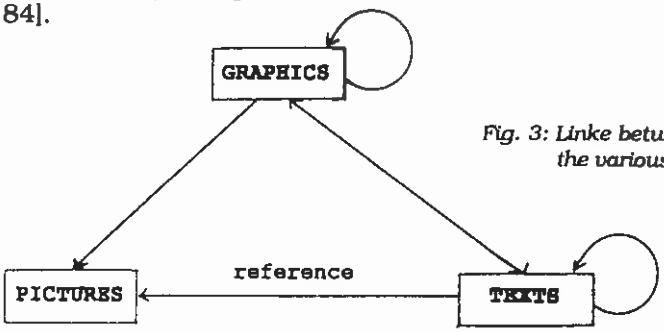


Fig. 3: Linke between the various types of base items.

The texts and graphics can refer to any other base items but the images sequences can only referred to.

Base organization can therefore be summed up with the following diagram:

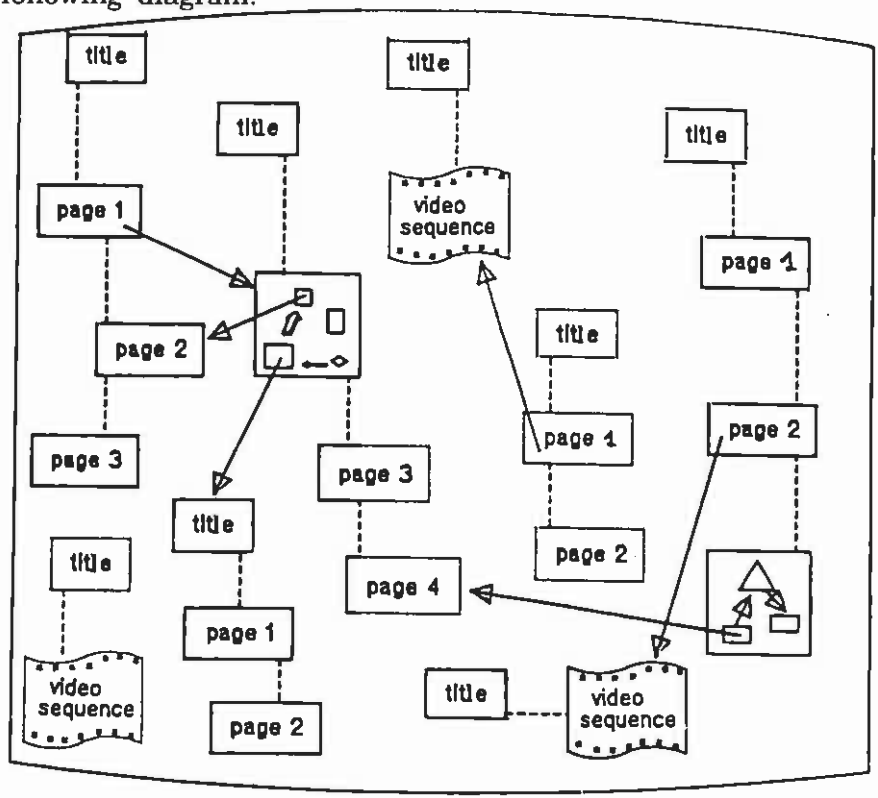


Fig. 4: Generalised Information Base.

IV - AUTOMATIC TITLE INDEXING

The main textual and visual descriptors of a card, are contained in its title. It is then sufficient to perform the indexing of the documents on the set of the base titles. Thus, when a new card is collected, the system analyzes the title to extract the significant (key) words and integrate them into a thesaurus.

The extraction procedure for keywords is as follows:

- delection of separators (comma, ...);
- delection of the words which do not convey any semantic meaning (articles, pronoms, ...)
- truncation of the remaining words down to 7 characters (parametric value usualy sufficient to differentiate two words) in order to optimize the thesaurus volume and the response times of the system. The results of a statistical study ([TUF 84]) justify the choice of a 7 characters radical and prove the steady performance of the thesaurus.

The set of keyword thus obtained from the title of a card is called reduced title.

Moreover, each keyword is assigned a calculated semantic weight, the value it will hold inside the thesaurus. This weight in function:

- of the rare occurence of the word in the totality of the base titles;
- of the way the word is "outlined" in the text (visualization attributes);
- and in some cases of the importance granted by the administrator. [ANT 84] [TUF 84].

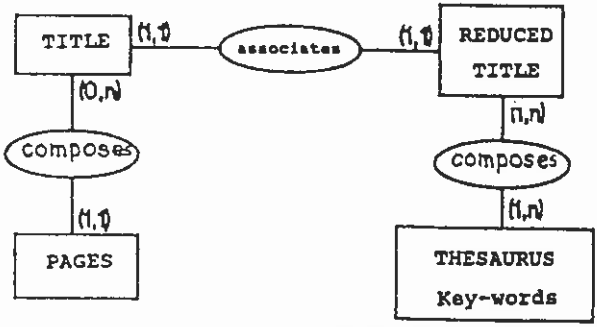


Fig. 5: Conceptual system schema (Entity/Relationship model).

In the same way, the administrator defines a picture thesaurus composed of key-pictures extracted from each video sequence in the base. He must then associate the key-pictures to the cards, which will allow the system to calculate the weight of each element in the picture thesaurus.

V - INFORMATION ACCESSING

It seemed important for ONISEP advisers that the AQUEDUCT system be easily accessed by all sorts of publics: adults as well as secondary school pupils. It is the reason why we have diversified interrogation modes (see Fig. 6, next page).

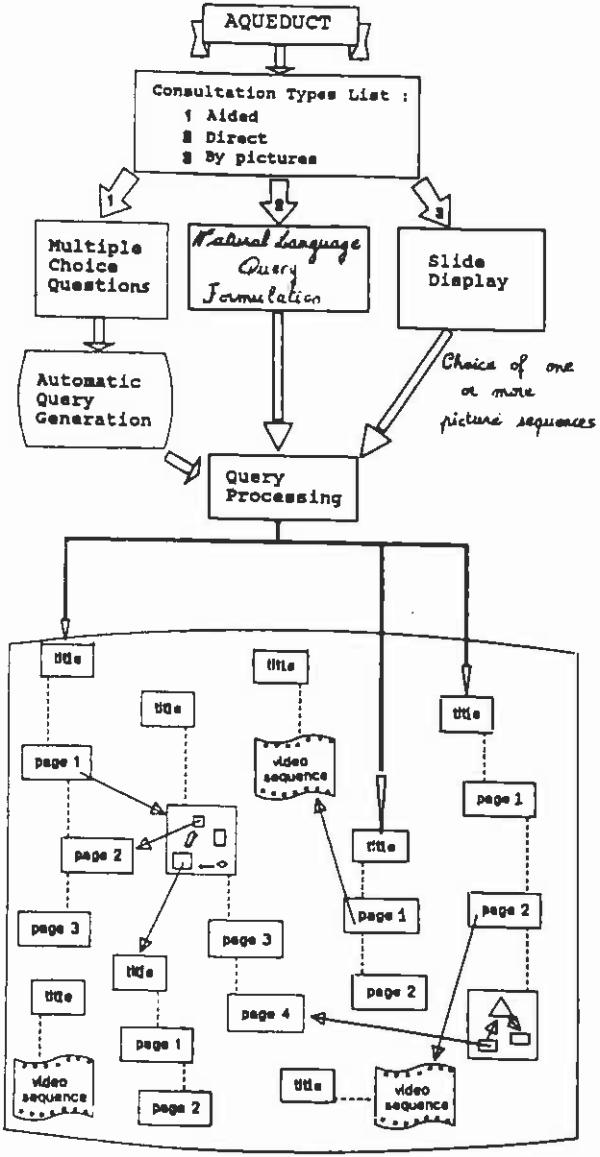


Fig. 6: AQUEDUCT whole consultation diagram.

A - QUASI-NATURAL LANGUAGE QUERY

The user need only express a natural language request. The system then analyses the quation as follows:

- creating a reduced text using the same process as when constituting a reduced title;
- searching through and extracting from thesaurus those titles that include at least one of the keywords of the reduced text;
- calculating the degree of similarity between the question and the selected titles. [TUF 84].

example: Let us consider the following request (see pictures in the Annex).
"I have just passed the 'BAC'. I would like to do trade and management"

The structural analysis of the request and the comparison with the thesaurus enhance the following keywords:

BAC TRADE MANGEMENT

B - AIDED CONSULTATION

This access mode is designed for users who find it difficult to formulate a request, through lack of vocabulary, accuracy or spelling mistakes. the consulting user (consulter) is then offered a multiple-choice questionnaire from which the system generates a request composed of key-words associated to each question. this request is then processed is the same way as in the previous consulting mode.

The ONISEP advisers mastery of interviewing techniques has allowed us to design this questionnaire. Each question is devised according to the user's previous answer thus making it possible to get a more accurate approach to his needs.

example:

Question 1: "How would you like to work:
outdoors, in an office, within a team ..."

Someone may answer:
Response 1: "in an office"

The system will then ask the following question:
Question 2: "What line would you like to be in:
accountancy, computer-science, ..."

The same person may answer:
Response 2: "computer-science"

The question asked by the system will be:
Question 3: "Which is the latest degree you have passed CAP, BEP, BAC, BTS, DUT, ..."

The user may answer:
Response 3: "Bac"

The request generated by the system will thus contain the following key-words:

"OFFICE COMPUTERS BAC"

C - CONSULTING THROUGH PICTURES

Using pictures as a means of access to information increases the liveliness of the dialogue with the consulter and gives the services a genuine multi-media dimension. Unfortunately problems related to picture-transport currently lead to two service-configurations:

- if the covered area is equipped with a fibreoptic-cabled network, the server will broadcoast pictures through the network, without any problem;
- when there is no cabled network, one must provide as many audio-visual sources (videodisks, videodisk readers or video cassettes) as consultations stations, which considerably increases the system implementation costs.

The videodisk is characterized by a large storage capacity (54.000 fixed pictures, 30 minutes' motion pictures), and a satisfactory information access speed (an average of 1 or 2 seconds). It has been chosen as an audio-visual information medium for the AQUEDUCT system.

AQUEDUCT picture access system requires the user to choose among a given number of pictograms, fixed or motion pictures, that are proposed to him or her from videodisks linked to the consultation station. [DAG 87]

These key-pictures are compared to those in the picture thesaurus. Their semantic weight allows the system to calculate the similarity degree between the request and the base documents. The consulter can then visualize the cards displayed in decreasing similarity degree.

VI - HARDWARE CONFIGURATIONS

A - MULTIMEDIA SERVER

AQUEDUCT is currently implemented on a Videotex server. Using the Public Telephone Network (PTN) has thus led us to set up the picture support on the actual consultation spot.

The resulting hardware configuration may represented as follows:

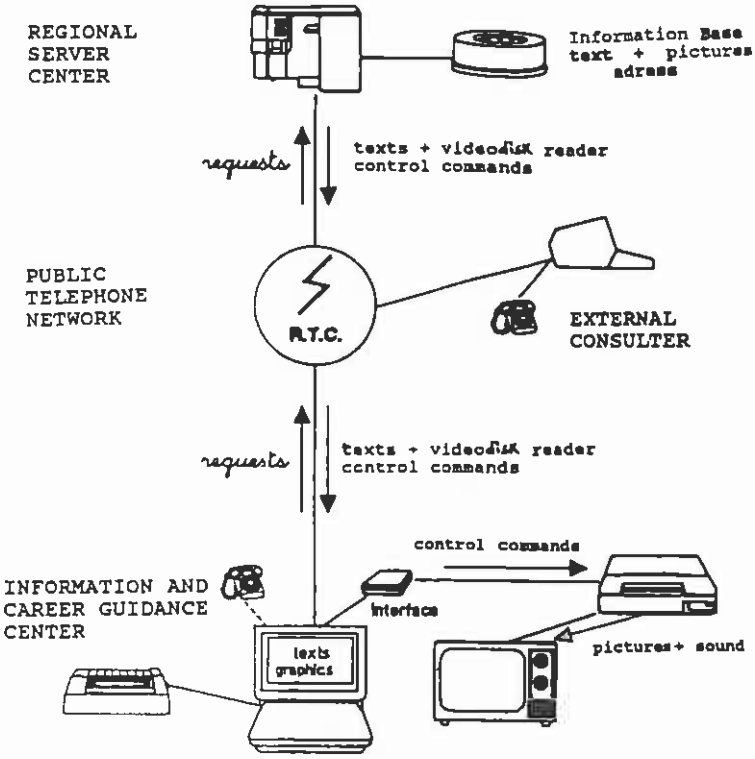


Fig. 7: Multimedia system hardware configuration.

Thus, when consulting from a privileged spot (CIO) supplied with the adequate audio-visual equipment:

- the user formulates, his or her request in one of the three available access modes;
- the request is dispatched to the server through the PNT and eventually processed by AQUEDUCT.
- the server displays the text on the consultation screen and triggers the videodisk reader.

B - AUTONOMOUS CONSULTATION STATION

Autonomous consultation stations, or multimedia stations, are widely available to the public (in townhalls, secondary schools, railway stations, ...) in order to reach as many people as possible.

Access to information can be achieved only "picture" mode, to make system consultation more lively and attractive in such passing places.

In addition to the information described in the previous paragraphs, the system offers the consulter a series of evaluation games that allow him to test his understanding of the received information.

On the other hand, the career guidance adviser is provided with an author-system that allows him or her to creat and act on the consultation graph structure, with reference to the users or the message to be conveyed.

An autonomous consultation station consists in:

- a micro-computer (CPU only);
- a two-entry television monitor (avideo entry, and a RGB one);
- a videodisk reader;
- a simple keyboard.

These components are assembled and sealed in an anti-theft box. The monitor screen alone is visible. The keyboard itself is protected by a 2 cm thick pane.

Research is now being developed to achieve base updating by teleloading, from a videotex-server..

VII - CONCLUSION

A - STATE OF ADVANCEMENT

The AQUEDUCT system is currently running on the BULL DPS-8 server of the inter-region computer center in Toulouse. The "text" component of the jobs and professions base is already operational and can be queried is natural language from videotex standard terminals. In the medium term, the documentary base will contain more then 300 professions.

By way of experiment, 20 CIOs and 20 secondary schools (lycées in Aquitaine are considered as pilot sites so avaluate the impact of such a system on a young public and show up which improvements ought to be achieved. Consultation statistics allow to analyze the main reasons for failures (problems of synonyms and mispelling). A spelling corrector is being added to the system in order to cope with the consequences of the numerous typing mistakes. (VORTEX system, CERFIA Laboratory).

Concurrently, the "picture" component is developed on an autonomous consultation station. The control a telematics server through the TELETEL network, which is essential to ensure a good "text-picture" integration, is being completed.14 videodisks are planned for the description of the 100 professions already available.

The practical application of the "graphics" component is not so advanced. It is actually dependent on technological choices, such as graphic screen definition and CD/ROM utilization for local storage.

B - PRESENTATION OF APPLICATIONS

In the course of working sessions, we will present a query model of the base made of:

- a terminal complying with videotex standards (Minitel);
- a videodisk reader connected to the Minitel.

This presentation will be made in connection with the CICT server (if the liaison with the French TELETEL network is technically possible).

BIBLIOGRAPHY

[AND 83] ANDRE J.
"Actes des journées, sur la manipulation des documents, Rennes 1983".
INRIA - 1983

[ANT 85] ANTON J. P. - DAGORRET F. - TUFFERY M.
"Um système d'aide à la recherche documentaire aux normes vidéotex:.
Rapport interne au laboratoire CERFIA.
Toulouse - 1985.

[ANT 87] ANTON J. P.
"Contribution au développment de systèmes Vidéotex Multimédia".
Thèse d'état, Université Paul Sabatier - Juillet 87

[BOR 86] BORNES C.
"Les nouveaux outils du spécialiste de l'information".
Support de cours INRIA - 1986

[CHR 85-a] CHRISMENT C. - ZURFLUH G.
"Advanced Data Bases. Multi-media Interfaces.
Principes de la communication Home-Machine: Parole, Vision".
Séminaires INRIA - Versailles - 1985

[CHR 85-B] CHRISMENT C. - CRAMPES J. B. - ZURFLUH G.
"Bases d'informations généralisées".
Monographies AFCET DUNOD - 1985

[CON 85] CONCHON A.
"Projet CONCERTO: le poste de travail".
L'Echo des recherches - CNET Paris A - 1985

[DAG 87] DAGORRET F. - LARRIEU F.
"Une base d'informations sur les métiers et les enseignements pour les lycées d'Aquitaine: le système AQUEDUCT".
Convention Informatique Latine - Barcelone - 1987

[QUI 86] QUINT V. - VATTON I. - BEDOR H.
"GRIF: An Interactive System for Structured Document Manipulation".
Nottingham - Avril 1986.

[TUF 84] TUFFERY M.
"System documentary, base de données textuelle - Le projet ETOILE".
Thèse de 3ème cycle - Université Paul Sabatier.
Toulouse - 1984

[VEL 84] VELEZ F.
"Un modèle et un langage pour les bases de donées généralisées: Projet TIGRE".
Thèse de Docteur Ingénieur - INP Grenoble 1984.