

Fenêtres ouvertes sur la connaissance

Dans le programme Ingénierie de la connaissance et des images, onze laboratoires numérisent les ressources documentaires de l'Université pour créer des outils pédagogiques innovants

Par Sandrine Lopez Photos Claude Pauquet

Programme transversal, inscrit au 12^e contrat de plan Etat-Région, l'ingénierie de la connaissance et des images (ICI) implique onze laboratoires¹ de l'Université, du Futuroscope et de l'Ensm, qui seront fédérés pour constituer la base d'un pôle de recherche à la fois fondamentale et appliquée.

Il s'agit de numériser les ressources documentaires de l'Université. Pour cela, les chercheurs s'intéressent aux contenus, à la numérisation en tant que telle et à la navigation dans ces contenus numérisés. A la clé, des logiciels et applications à inventer ou à adapter à l'objet numérique (site web ou CD-Rom) mais aussi des matériels nouveaux à acquérir (les scanners du commerce sont inutilisables pour des documents picturaux anciens), la diffusion au grand public d'un fonds considérable de documents.

Les laboratoires travaillent en parallèle selon deux axes de recherche proches de leur discipline initiale. D'une part, les linguistes et les littéraires, les psychologues et les informaticiens étudient la «Cognition : environnements documentaires interactifs». D'autre part, les laboratoires de l'Ensm se concentrent sur l'«Image numérique : simulateurs et environnements virtuels». Ce second axe de recherche est développé depuis plusieurs années à l'Ensm (voir *L'Actualité* n° 43).

«LA TRANSVERSALITÉ NE SE DÉCRÈTE PAS, ELLE SE CONSTATE»

L'étude des environnements documentaires interactifs (EDI) constitue le noyau dur du programme ICI. «*Elle justifie la transversalité*», selon Daniel Gaonac'h, directeur du LaCo. Ce projet de recherche est, en outre, plus récent même si Daniel Gaonac'h considère qu'il est «*la suite logique de relations nouées entre universitaires et, à l'extérieur, avec les professionnels, depuis cinq ou dix ans. La transversalité ne se décrète pas, elle se constate. Mais, il est vrai que la création*

de la Maison des sciences de l'homme et de la société facilite ce genre de collaboration.»

Il insiste sur le rôle modeste du LaCo dans ce programme mais concède que ce laboratoire joue souvent un rôle d'interface entre les scientifiques et les littéraires. Car les liens entre tous sont loin d'être évidents. «*On ne sait même pas encore si certains se concrétiseront*, ajoute-t-il. *Par exemple, le travail de Michel Léard, à l'Ircm, sur la visualisation scientifique pour l'imagerie médicale est très spécifique. Notre idée, c'est qu'il doit pouvoir rejoindre la visualisation des documents littéraires anciens.*»

L'apport de chaque laboratoire doit apparaître à différents niveaux dans les recherches. Le projet EDI regroupe trois thématiques qu'il faut jumeler, au fur et à mesure des résultats. Première thématique : les contenus. Très nombreux et divers. «*Il ne faut pas confondre : numériser les documents n'est pas l'objectif de ce programme mais son moyen. Numériser pour numériser ne nous intéresse pas. L'objectif est de réfléchir aux nouveaux traitements des documents numérisés.*»

NUMÉRISER POUR VOIR L'INVISIBLE

La seconde thématique porte sur la numérisation même. Informaticiens et linguistes vont devoir trouver les procédures les plus efficaces pour analyser ces contenus. Comment restituer les accents étrangers des documents sonores ? Doit-on tenir compte de toutes les nuances de couleurs ? Ou encore, les travaux de Jean-Louis Duchet, professeur d'anglais et responsable du Forell, qui portent sur des manuscrits anciens. Un même caractère imprimé avec du plomb comporte toujours des petites différences. Le dialogue entre le linguiste et l'informaticien devra définir la marge de reconnaissance de ces caractères «plombés».

Troisième thématique, qui doit être intégrée dans les deux premières : la navigation dans les documents



Daniel Gaonac'h

numérisés. Quelle ergonomie sera la plus efficace ? *«Il s'agit de problèmes purement cognitifs, poursuit Daniel Gaonac'h. Nous cherchons à connaître les réactions de l'être humain devant un ordinateur. Comment faire en sorte que l'esprit humain puisse utiliser des informations dispersées dans un système numérique, sans qu'elles soient immédiatement apparentes ? Comment faire aussi pour distinguer des informations qui n'ont pas toutes la même source : voyez par exemple les difficultés qu'on a parfois à distinguer l'information de la publicité sur Internet.»*

La numérisation aura pour effet immédiat de rendre une masse considérable de documents accessibles au grand public. Certains de ces documents ne sont actuellement accessibles qu'aux chercheurs. Ces derniers ajouteront aux documents les études qui s'y rapportent. Ils tiennent aussi à identifier clairement les sources. Ainsi, la numérisation permettra de tracer les chemins qui mènent à l'origine d'un document, d'un texte, voire d'une phrase et d'un mot. *«Nous pensons que la numérisation peut permettre, par exemple sur le plan littéraire, de découvrir des choses que l'on ne verrait pas autrement»*, précise Daniel Gaonac'h.

Les nouvelles technologies décuplent les capacités de diffusion des informations. Et les chances de se perdre dans Internet. Mais à terme, l'homme devant la machine ne doit plus être un consultant. Il pourra «interagir» et à son tour analyser les documents, les critiquer, ajouter ses propres notes. ■

1. Pour la partie "Environnements documentaires interactifs" : le Centre de recherches latino-américaines et équipe Archivos, le Forell (Formes et représentation en linguistique et littérature), le Centre d'études supérieures en civilisation médiévale, le laboratoire langage et cognition, Migrinter (migrations internationales), le laboratoire d'informatique scientifique et industrielle (Lisi), l'institut de recherche en communication et micro-ondes - équipe Signal, Image et Communication (Ircom-Sic).

Pour la partie, "Simulateurs et environnements virtuels" : l'Ircom-Sic, le Lisi, le laboratoire de mécanique des solides, le laboratoire d'études aérodynamiques, le laboratoire d'études thermiques, le laboratoire de métallurgie physique.

Une nouvelle pédagogie

Depuis une dizaine d'années, le LaCo étudie les stratégies individuelles de recherche d'informations dans un système hypermédia complexe. Ces études ont abouti à la fabrication de programmes hypertextes simples sur lesquels Jean-François Rouet teste actuellement les aptitudes des étudiants. Ces programmes trouveront une application directe dans le projet EDI. Dans un box totalement blanc, les étudiants se retrouvent face à un ordinateur qui leur pose une question d'ordre général : quelle a été la consommation d'eau domestique, pour l'année 1997 en France ? *«On enregistre tout ce que fait la personne pour trouver la réponse, explique-t-il. Quelle partie du document elle consulte, si elle trouve directement ou fait des détours, etc. Il s'agit de constater si l'organisation du document a une influence sur la démarche intellectuelle.»*

Le document est basique au regard de ce que l'on trouve sur CD-Rom ou Internet, mais il permet d'étudier les processus mentaux qui se déroulent dans ce genre de tâches. Le LaCo espère ainsi mettre en évidence les interfaces informatiques qui maximiseront la recherche et l'exploitation d'informations dans un système multimédia. Jean-François Rouet travaille en relation avec des concepteurs de produits informatiques qui, depuis quelque temps, s'intéressent à la psychologie des utilisateurs. *«On est très vite désorienté dans un produit multimédia, indique-t-il. Aujourd'hui, les industriels et les fabricants viennent aux sciences humaines parce que nous*

avons mené dix ans d'expériences, notamment avec des enfants, qui tendent à prouver que le multimédia ne réduit ni les différences sociales, ni les différences culturelles. Les produits très sophistiqués ont même l'effet inverse : ils n'aident que les meilleurs.»

Pour autant, l'intérêt pédagogique des outils multimédia est *«indéniable, dans la mesure où l'on suit la logique des étudiants dans l'interaction compréhension/production d'informations»*.

Le LaCo a, par ailleurs, travaillé sur l'élaboration d'un CD-Rom pour le Cript, centre de développement multimédia créé par un réseau d'écoles d'ingénieurs en télécommunication. *«Le Cript est chargé de développer des supports pédagogiques innovants. L'objectif est d'aider les élèves, qui ont un savoir trop scolaire, à réfléchir dans un sens plus pratique afin qu'ils s'insèrent plus rapidement dans des projets d'entreprises.»*

Le CD-Rom, intitulé *L'électronique en question*, a été présenté au salon de l'Education. Il va maintenant être utilisé par les élèves ingénieurs.



Développer les partenariats avec les grands organismes de recherche

Sur les 43 laboratoires reconnus par le ministère de l'Enseignement supérieur et de la Recherche que comptent l'Université de Poitiers et l'Ensm, 24 sont associés au CNRS et constitue une unité propre de recherche du CNRS. Le MSHS est une unité mixte de services (qui associe de nombreuses équipes de recherche). Des laboratoires travaillent avec l'Inra, à Lusignan (en biologie végétale) et Surgères (transplantations multi-viscérales). De plus, il existe un objectif extrêmement fort d'association avec l'Inserm, envisagé d'ici quatre ans au maximum. L'intérêt de ces diverses accréditations se situe à plusieurs niveaux. D'un point de vue de l'expertise tout d'abord, c'est un gage de qualité,

assurant aux laboratoires une reconnaissance nationale et internationale (grâce à une procédure normalisée d'évaluation). C'est ensuite un appui financier et un apport en personnel technique et de recherche importants. *«Il faut souligner – et cela est primordial – que le CNRS participe, en étroite collaboration avec l'Université, à la formation continue des personnels et qu'il est partie intégrante de la cellule de valorisation de la recherche (commune à l'Ensm et l'Université), et de l'incubateur Poitou-Charentes»*, précise Jean-Paul Bonnet, directeur de recherche au CNRS (à l'Ensm), vice-président de l'Université chargé du conseil scientifique.

Les Bénédictins numériques

Les applications du projet "Environnements des documents interactifs" sont multiples et de deux ordres. La numérisation à des fins d'édition et de diffusion offrira, d'ici quelque temps, la possibilité à chacun de se faire fabriquer un livre à la demande. Vous surfez sur le web, trouvez le livre précieux que vous cherchiez depuis des lustres, pianotez les numéros de votre carte de crédit, votre adresse... Le lendemain, le livre "introuvable" est dans votre boîte à lettres, relié et broché. Irréaliste ? Pas selon les chercheurs de l'Université qui imaginent mal comment une maison d'édition pourrait se passer de cette valorisation de sa collection.

Mais c'est la numérisation à des fins de constitution de corpus littéraires, historiques et linguistiques qui intéresse particulièrement les chercheurs poitevins. Elle doit permettre de réaliser une analyse en profondeur des textes, proche de l'activité sur les textes sacrés des moines bénédictins. Mais «ce qui leur prenait toute une vie, pour annoter et analyser la Bible, se fait en quelques minutes avec l'ordinateur», indique Jean-Louis Duchet du Forell. Miraculeux, non ?

Il voit dans la numérisation un autre avantage: «Certains documents du XVI^e au XVII^e siècles n'ont pas été réimprimés depuis l'origine et vont pouvoir revivre numériquement. On ne travaillera plus avec 500 ou 600 grands textes connus mais avec 3 000 ou 5 000 ! Il y en aura sans doute de valeur inégale, mais cela donne une profondeur de champ dans l'étude sans commune mesure avec ce dont on disposait jusqu'à présent. Le travail d'investigation des chercheurs sur le texte et la pensée de la Renaissance et de

l'époque moderne va changer et trouver de nouveaux prolongements.» Le travail de numérisation conduit en ce sens par Marie-Luce Demonet complètera les collections numérisées cette année par la Bibliothèque nationale de France.

Jean-Louis Duchet travaille sur la représentation de la prononciation de l'anglais dans les dictionnaires anciens. Son collègue Philippe Caron étudie, pour sa part, les dictionnaires français. Ils utilisent chacun leur technique, qui a pour point commun d'être pour l'instant "artisanale". Jean-Louis Duchet passe par la reconnaissance optique des caractères, avec des scanners et des logiciels de bureautique. Philippe Caron fait saisir ses dictionnaires par un claviste. Mais, selon le souhait du ministère de l'Education nationale, Poitiers pourrait devenir un centre de ressources technologiques dans le domaine de l'industrie des connaissances. Dans ce cadre, il est prévu d'acquérir des matériels qui accéléreront la numérisation et le balisage des textes, des images et des sons, en ayant recours au langage SGML, langage plus sophistiqué que l'actuel HTML qui est celui du web sur Internet.



Claude Pauquet

mshs

Un réseau de compétences à inventer

Association de compétences, interdisciplinarité. Ces termes reviennent souvent dans la bouche de Claire Gérard, directrice de la Maison des sciences de l'homme et de la société (MSHS, unité rattachée au CNRS), fondée en 1996 à l'Université de Poitiers.

Il existe douze maisons de ce type en France, certaines étant thématiques. Se greffent à elles des instituts de recherche qui fédèrent plusieurs laboratoires mais qui n'ont pas une mission de développement de programmes de recherche interdisciplinaires. «La première mission de la MSHS est de restructurer les sciences humaines, c'est-à-dire de modifier les anciens cloisonnements disciplinaires pour associer les compétences des sciences humaines et des sciences de l'ingénieur, par exemple», explique Claire Gérard. Il s'agit d'un changement notable dans le comportement des chercheurs, pour la plupart «heureux de sortir de leur discipline et de se mettre autour d'une table pour découvrir le voisin».

Héberger des labos, les soutenir et les dynamiser, conduit inévitablement à des projets communs de recherche. «Si nous avons ajouté le mot société à notre Maison des sciences de l'homme, c'est

pour que des chercheurs qui travaillent sur la pollution intègrent des géographes et des économistes à leur équipe. Tout problème social est interdisciplinaire», dit-elle en posant vite un bémol: «L'interdisciplinarité est un complément nécessaire mais ne doit pas diluer les spécialités, souvent très pointues.»

La MSHS a initié neuf programmes de recherche à Poitiers mais aussi avec la Maison des sciences de l'homme de La Rochelle.

Seconde phase: la mise en réseau. La MSHS de Poitiers représente les douze maisons au sein du comité national de pilotage, créé en octobre dernier, pour animer ce réseau.

«Outre les missions locales des maisons, il y a une volonté de collaborer entre les chercheurs de chaque MSH, indique Claire Gérard. La finalité est de créer un réseau d'experts, d'éditer des publications haut de gamme, d'inviter des chercheurs étrangers pour développer des programmes internationaux.»

Ce réseau devrait permettre à de jeunes chercheurs d'obtenir des bourses post-doctorat pour des programmes ou des spécialités susceptibles d'intéresser plusieurs laboratoires.

Claude Pauquet