



[www.ichim.org](http://www.ichim.org)

Les institutions culturelles et le numérique  
**Cultural institutions and digital technology**

École du Louvre  
8 - 12 septembre 2003

**MAVI : HISTOIRE ET PATRIMOINE CULTUREL DU  
MOYEN-ORIENT ANCIEN À L'ÉPOQUE PERSE**

**Pierre Briant – José Paumard – Marie-Françoise  
Clergeau**

« Acte publié avec le soutien de la Mission de la Recherche et  
de la Technologie du Ministère de la Culture et de la Communication »

## Résumé

Le projet MAVI (Musée achéménide virtuel et interactif) a pour objectif de rassembler et de disposer sur un site-internet tous les objets et images (sceaux, monnaies, stèles, peintures etc.), relevant d'un espace-temps historique bien identifié (le Moyen-Orient de l'Indus à la Méditerranée, sous la domination des Perses achéménides, c. 550-330 av. n.è.), et de participer ainsi directement à l'inventaire et à l'analyse d'un patrimoine culturel unique, actuellement dispersé entre des dizaines de musées et de collections dans le monde. Grâce à l'utilisation de procédures de visualisation interactives au travers de l'Internet, d'images numérisées en très haute résolution, les chercheurs pourront travailler en direct sur l'iconographie, dans des conditions de travail incomparables, puisqu'ils pourront non seulement avoir accès à une base iconographique et à des images de haute qualité, mais qu'ils pourront, en outre, avoir une action sur ces images par des commandes en ligne à leur disposition sur l'écran. Prenant appui sur les travaux scientifiques et en favorisant leur développement, ce Musée sera lié à une base de données iconographiques interrogeable en ligne (sous Oracle). Afin d'être en mesure de résoudre dans toute leur ampleur les problèmes scientifiques et techniques liés à une telle entreprise (sans équivalent au monde), nous avons lancé un vaste projet collaboratif, fondé sur l'inter-disciplinarité établie entre des spécialistes des Sciences humaines (historiens, archéologues, historiens de l'art), et des spécialistes de la consultation et du transport d'informations multimédia. Le projet est co-piloté par la «Chaire d'histoire et civilisation du monde achéménide et de l'empire d'Alexandre» du Collège de France (Prof. Pierre Briant) et le personnel permanent mis à disposition par le Collège de France (Marie-Françoise Clergeau, sous-directeur de laboratoire, Valérie Janicot, technicienne), et par le «Laboratoire de traitement et de transport de l'information, L2TI» (Prof. Ken Chen) représenté par José Paumard (Maître de Conférences). Il s'appuie sur l'expertise d'un comité éditorial international, et sur un réseau international de collègues et de conservateurs de musées.

## Définition du projet et origines

Le projet MAVI est né d'une analyse de la situation et du développement nouveau d'un projet, qui a été lancé en 2000 par Pierre Briant, depuis la chaire d'histoire et civilisation du monde achéménide qu'il occupe au Collège de France, avec l'aide du Ministère de l'Éducation Nationale et le Ministère de la recherche, le projet achemenet.com (<http://www.achemenet.com>). Le projet MAVI partage avec ce premier projet un constat de départ, qu'il convient de rappeler, même brièvement. Achemenet a pour but de rassembler sur un site web tous les documents primaires relatifs à l'histoire d'un espace-temps que l'on peut ainsi définir: de l'Indus aux Balkans, et de l'Asie Centrale à Assouan, –de la conquête et de l'unification politique de cet espace sous Cyrus le Grand (à partir de c. 550 av. n.è.) jusqu'à l'éclatement des structures unitaires, après la mort d'Alexandre le Grand (soit vers 300).

Achemenet.com, qui a été particulièrement bien accueilli<sup>1</sup>, est destiné en premier lieu aux chercheurs, qui sont à la fois acteurs, réalisateurs et utilisateurs. Ils sont très dispersés à la fois entre de nombreuses aires culturelles antiques, et entre de nombreux pays actuels. Selon un paradoxe qui n'est qu'apparent, la création d'un tel site Internet a donc permis de passer d'une communauté scientifique virtuelle à une communauté scientifique réelle. Sur les attendus et les objectifs du projet, voir l'Appel à collaboration disponible en ligne à <http://www.achemenet.com/pdf/appel.pdf> et <http://www.achemenet.com/pdf/call.pdf>. A l'heure actuelle, le projet achemenet.com continue son enrichissement, grâce à la collaboration active de nombreux collègues de par le monde et dans de nombreuses spécialités. Il est actuellement répertorié comme le «[Portail des études achéménides](#)» sur des dizaine de sites spécialisés dans de très nombreux pays.

Pierre Briant a eu plusieurs fois l'occasion de présenter ce projet dans des colloques ou lors de conférences spécialisées en France et à l'étranger, depuis la première présentation de la maquette le 12 juillet 2000 lors de la Rencontre Assyriologique Internationale (Paris, Collège de France), et les colloques spécialisés réunis au Collège de France en décembre 2000 et en juin 2001 (voir compte rendu dans *Annuaire du Collège de France 2000-2001*,

---

<sup>1</sup> Voir *L'Histoire*, Octobre 2000, p. 106 [Historiens et Géographes](#), janvier-février 2002. L'excellence du site a été soulignée par le Ministre de l'Éducation Nationale lors de sa conférence de presse du 14 janvier 2002 sur «[Les arts et la culture dans l'enseignement supérieur](#)». Sur les quatre derniers mois, le site a reçu entre 2600 et 3200 visites/mois.

p. 725-729). Par exemple à Téhéran, lors du Colloque sur «Les recherches archéologiques françaises en Iran» (octobre 2002), il a présenté une communication sur «Remarques sur l'évolution récente des recherches achéménides et propositions pour une collaboration internationale». Par ailleurs, à l'occasion du Congrès du groupe des Bibliothèques d'Art réuni à Bordeaux (14 mars 2002), une communication a été présentée par P. Briant, J. Paumard et M.F. Clergeau sur «Vers la création d'un musée achéménide virtuel et interactif» (Actes sous-presses). A Chicago (Oriental Institute), une réunion du projet MAVI a eu lieu les 11-13 avril 2002 avec nos collègues américains, en présence de P. Briant et de M.F. Clergeau.

## Contexte et enjeux de société

L'enregistrement, l'archivage et la consultation d'un patrimoine culturel sont aujourd'hui et pour demain des enjeux décisifs du développement de la vie culturelle et scientifique. Le début du XXI<sup>e</sup> siècle représente un moment particulier où les conditions permettent d'envisager des solutions techniques à des problèmes auxquels l'existence des seuls musées «réels» ne peut répondre : sauvegarder et rendre accessible ce patrimoine. Il y a un siècle, l'on peut imaginer que le projet aurait été celui d'un «Corpus des documents iconographiques liés à la civilisation achéménide» en douze volumes in-folio : la réalisation se serait étendue sur un siècle, et la mise à jour en aurait été rendue quasiment impossible. Quant à l'accès des chercheurs et du public, il n'aurait été possible que dans quelques institutions privilégiées capables d'acquérir des ouvrages hors de prix. Aujourd'hui les progrès joints de la numérisation, des bases de données et des transferts via Internet rendent possible ce qui ne l'était pas hier. Il est donc de la responsabilité des chercheurs et des institutions de recherches de faire aujourd'hui ce travail de rassemblement, d'archivages et de mise à disposition de ces immenses archives artistiques, archéologiques et culturelles, aujourd'hui encore dispersées entre des centaines de lieux et de publications, musées et réserves, catalogues, rapports de fouilles, articles et ouvrages. Il est de notre responsabilité sociale d'impulser un vaste projet international, où collaboreront non seulement les spécialistes du champ considéré (historiens, archéologues, conservateurs de musées), mais aussi les spécialistes venus des sciences humaines et sociales, et spécialistes venus des secteurs des sciences et techniques, capables de mettre au point, en échange intime avec les premiers, les

instruments et procédures susceptibles de rendre le projet réalisable et accessible à tous, dans tous les pays.

## **Travaux existants sur le même sujet en France ou à l'étranger**

Dans le domaine considéré, -- l'histoire et l'archéologie des pays du Moyen-Orient ancien, -- plusieurs projets sont nés ici et là, singulièrement aux Etats-Unis, mais aussi en Angleterre et en Espagne, mais la plupart ont avorté au bout de quelques mois ou de quelques années, en raison d'un objet parfois insuffisamment identifié, en raison aussi d'une inadéquation entre le projet et les moyens de recherches. A l'heure actuelle, un projet s'est lancé aux Etats-Unis, la «[Uniform Digital Library Initiative](http://cdli.ucla.edu/)» (University California Los Angeles et Max Planck Institute <http://cdli.ucla.edu/>), qui vise à rassembler sur un site-web l'ensemble des tablettes sumériennes. En revanche, il n'existe pas, au monde, une seule entreprise connue, qui soit explicitement vouée à la collecte et à l'archivage exhaustifs, et à la publication en ligne et à l'étude d'un patrimoine culturel aussi précisément désigné que celui qu'entend intégrer le MAVI.

Il existe bien entendu des sites organisés par des musées. Mais notre projet est complètement différent

- sur les sites des musées, les images sont fixes, quelque fois disponibles en deux résolutions dont la plus fine reste insuffisante pour une réelle exploitation scientifique.
- le musée ne met pas en ligne toutes ses images il s'agit d'une sélection opérée non en fonction de critères de la recherche mais en fonction d'une appréciation sur l'esthétique de l'image en question (pour ne pas parler des objectifs commerciaux).
- il est une autre différence, plus fondamentale encore. Les Musées, qui disposent de nombreuses collections hétérogènes (vases grecs, peinture flamande, tapisseries des Gobelins etc.) tentent de donner une image de cette diversité, selon une logique de communication, rendue plus efficace par l'utilisation d'un moteur de recherches (dans le meilleur des cas).

Tout au contraire, la logique du projet MAVI est à la fois scientifique, thématique et interactive. Un musée virtuel thématique est une très grande nouveauté. Il sera organisé autour d'un projet scientifique extrêmement cohérent, --il servira à la communauté scientifique internationale, --il sera le résultat d'une collaboration institutionnelle étroite entre deux secteurs de la recherche, dans le domaine de l'histoire, de l'archéologie et de

l'iconographie d'une part, et, d'autre part, dans celui traitement des images et des réseaux informatiques et la consultation et transport des informations multimedia.

L'originalité du projet se mesure aussi au degré de collaboration très étroite entre deux secteurs de recherche qui trop souvent s'ignorent ou se méconnaissent : les Sciences humaines et les recherches dans le domaine de l'informatique et des transferts de données. Le rapport avec l'intitulé de la chaire de Pierre Briant est suffisamment explicite pour ne pas nécessiter de commentaire. Le projet n'utilisera pas les moyens de diffusion habituels des travaux de la chaire (cours et séminaires du Collège de France); il trouvera son expression sur Internet par la synthèse de procédures spécifiques. Dans le même temps, il répond de façon originale au dessein du Collège de France : "enseigner la recherche en train de se faire". C'est la première réalisation en vraie grandeur d'un de ces grands outils informatiques internationaux souhaités par les pionniers de l'utilisation de l'informatique dans les sciences historiques au Collège de France. La synergie est également évidente avec le GDR 2538 («Réseau international d'études et de recherches achéménides»), en particulier avec l'un des axes de recherches définis, «Les images de l'empire». Mais la différence de nature, d'échelle et d'objectifs n'en est pas moins très marquée, en raison principalement de la collaboration étroite avec les sciences et techniques, qui fonde le projet MAVI lui-même.

Le Laboratoire de Traitement et Transport de l'Information (L2TI, EA3043) est une entité de recherche créée par l'université Paris 13 en janvier 1998 et reconnue par le Ministère de l'Éducation Nationale, de la Recherche et de la Technologie (MENRT) comme une Équipe d'Accueil (EA3043) depuis la rentrée 1999. Rattaché à l'institut Galilée depuis sa création, le laboratoire compte dix-sept personnes dont douze enseignants-chercheurs titulaires répartis sur l'institut Galilée et sur l'IUT de Villetaneuse.

Les activités de recherche du L2TI s'inscrivent dans l'axe "Science et technologie" retenu par l'université Paris 13 comme l'un des axes prioritaires, plus particulièrement dans la subdivision "Sciences et technologies de l'information et de la communication" rubriques "Télécommunications" et "Systèmes d'information, de communication, multimédia". Les activités de recherche du L2TI couvrent deux secteurs : le traitement des images et les réseaux de télécommunications, avec un positionnement volontariste sur des thématiques émergentes résultant de l'interaction entre ces deux secteurs. Cette interaction se manifeste avec la généralisation de l'utilisation de l'Internet.

## Objectifs scientifiques et méthodologie.

Notre politique scientifique pourrait se résumer ainsi :

- Se positionner sur les thématiques émergentes telles que le transport des informations multimédia en vue de leur consultation en ligne, en particulier des images. Cette application est pour nous terrain d'application et une source d'inspiration.
- Développer des recherches sur les problèmes récurrents et des techniques de base, et les valoriser à travers des applications diversifiées.

Situé en synergie étroite avec le programme achemenet.com, le projet MAVI partage une méthode□l'utilisation d'une plate-forme Internet, mais il est en même temps spécifique et original. Du côté des historiens, archéologues et historiens de l'art, le programme scientifique porté par le MAVI s'intègre dans une réflexion très active depuis quelques années, en particulier dans le cadre d'un empire multi-ethnique, multilingue et multiculturel comme l'était l'empire achéménide□ Lancée particulièrement par deux collaborateurs du MAVI (Margaret Root, Ann Arbor et Mark Garrison Austin, Texas) et illustrée par leurs travaux (cf. leur premier volume des sceaux de Persépolis, Chicago 2002), ces recherches s'organisent de manière privilégiée autour de la question des rencontres entre traditions iconographiques et des influences réciproques centre/périphérie (cf. analyse des recherches 1996-2000 dans P.Briant, *Bulletin d'histoire achéménide II*, Paris, 2002, p. 191-206□ voir aussi l'ouvrage édité par Ch. Uehlinger, *Images as media. Sources for the cultural history of the Near East and the Eastern Mediterranean, Ist millenium B.C.*, Fribourg-Göttingen, 2000). Les chercheurs ont un besoin urgent d'avoir accès à un outil de consultation et de visualisation, tel que celui que propose le MAVI□

L'on peut considérer que le projet sera mené à bien à bien grâce à l'intervention conjointe et coordonnée de quatre entités□ trois entités scientifiques (Collège de France, CNRS et L2TI) et une entité de service.

La mise au point du projet MAVI a été préparée au cours de réunions (comme celle de Chicago en avril 2002), par des discussions très nombreuses entre P. Briant, J.Paumard, M.F. Clergeau et des collègues étrangers, et par un grand nombre de contacts, en particulier avec des Musées□

- La cellule du Collège de France (Pierre Briant, Marie-Françoise Clergeau, Valérie Janicot) est plus particulièrement chargée de rassembler la documentation sous forme essentiellement de textes et d'images, numérisées ou non. Elle procède également au traitement des images en vue de leur mise en ligne. Elle assure enfin l'identification des images, et l'enregistrement des données dans la base. Pierre Briant se charge plus particulièrement des contacts. Concernant la mise au point des normes et protocoles liés à la base de données, Marie-Françoise Clergeau travaille en très étroite collaboration avec José Paumard. Par ailleurs, elle supervise les travaux de scanning et d'enregistrement des données, qui sont effectués par une technicienne rattachée à la chaire (qui a reçu une formation spécifique), Valérie Janicot. Au sein de cette cellule, Margaret Root et Mark Garrison sont chargés plus spécifiquement d'une catégorie d'objets et d'images, les sceaux et empreintes de sceaux (rassemblement, publication, données) □ le développement de la cellule « Monnaies » est supervisé par Andy Meadows (British Museum) et par Koray Konuk (CNRS/GDR 2538 □ Comité de pilotage).

- La cellule du L2TI (José Paumard) met au point les procédés de visualisation interactifs sur Internet et construit et gère la base de données, également installe et gère le moteur de recherche extérieur.

- Pour chaque sous-ensemble iconographique, l'intervention de l'unité de services (société privée extérieure □ Philippe Bertin) se fait en continu, et en collaboration très étroite avec José Paumard, lui-même en charge de la mise au point des processus scientifiques et techniques et de leurs test et validation en ligne. Cette société extérieure (qui travaille au terme d'un contrat spécifique passé avec le programme MAVI) est chargée de la charte graphique (en association étroite avec P. Briant) et du graphisme, ainsi que de la fabrication des écrans et de leur téléchargement sur le site de l'IN2P3 (sous la direction technique et scientifique de José Paumard).

Le développement de ce projet s'articule autour de deux grands axes □

- Un axe technique □ le développement et l'intégration des outils de mise en place et de consultation de la base de données □

- Un axe documentaire □ le renseignement de cette base de données.

Ces deux axes progressent en parallèle, et certaines avancées de l'un sont nécessaires pour les avancées sur l'autre.

Techniquement, le projet MAVI se compose de trois modules principaux □



- Une base de données, dont la structure intègre les champs qui ont obtenu le consensus des différents spécialistes lors des réunions de juin 2001 à Paris, d'avril 2002 à Chicago, et de décembre 2002 et janvier 2003 à Londres.
- Une interface-auteur, qui permet d'administrer cette base de données, et qui permet aux indexeurs d'enregistrer des documents dans cette base.
- Une interface de consultation, au travers de laquelle sont présentés les documents multimédias. Elle comprend un certain nombre d'originalités, dont la présentation d'images de très haute définition, outil indispensable aux chercheurs qui veulent exploiter la base pour leur travail.

Ces trois modules sont des développements techniques sur lesquels le L2TI est maître d'œuvre, apportant ses compétences dans ce domaine. Nous revenons dans la suite sur les spécifications fonctionnelles de ce système. Elles ont déjà fait l'objet de nombreuses réunions et discussions, entre les futurs utilisateurs de MAVI d'une part, et le L2TI d'autre part. Le système est actuellement en phase de développement. L'objectif est de disposer d'une version beta à la fin de l'année 2003.

Le L2TI a également mis à l'étude une partie prospective portant sur les façons d'organiser, d'analyser et d'extraire de l'information des données de la base, ainsi que le développement des moyens d'accès et de navigations originaux.

## **Spécifications du système**

Rédiger des spécifications précises pour un tel système nécessite de faire la synthèse entre de nombreuses contraintes très diverses.

Le projet MAVI a été lancé à la suite du projet achemenet.com, et a pu bénéficier de toute l'expérience acquise lors de son développement. De nombreuses réunions, locales et internationales ont eu lieu autour de ce projet, ce qui a permis d'obtenir énormément d'informations sur les attentes des chercheurs, tant sur le contenu de la base, que sur les fonctionnalités que doit offrir son système de gestion.

Par ailleurs, nous nous sommes attachés à garantir un certain nombre de points qui nous semblaient importants pour l'exploitation scientifique de cette base, tels que l'intégrité et la fiabilité des données, le respect des droits de propriété en particulier pour les images, la traçabilité de l'information, la pérennité de la base.

Passons en revue les points les plus importants.

Les données que MAVI doit gérer sont de trois types□

- Des données textuelles, typées ou non (dates, lieux, ...), le plus souvent multilingues.
- Des images, de tailles, qualités et formats très divers.
- Des structures associant données textuelles entre elles, avec ou sans images. Ces structures doivent pouvoir se hiérarchiser aisément.

Les données enregistrées sont des images et des documents scientifiques. Ils arrivent au gré des collaborations internationales, à un rythme variable. MAVI doit donc être un système suffisamment dynamique pour accepter les documents de cette façon, et doit être capable de se réorganiser en fonction du contenu de ces documents.

Les collaborations en cours nous permettent de prévoir que quelques dizaines de milliers de documents devraient être enregistrés dans MAVI d'ici à deux ans. Il est difficile d'avoir une estimation fiable au-delà.

Les documents se regroupent naturellement en corpus. Les corpus doivent pouvoir être créés et modifiés dynamiquement. Cette structure est nécessaire aux chercheurs, et leur permet de créer les grandes catégories de documents.

Chaque corpus contient un jeu de champs qui permettent de qualifier les objets qu'ils contient. Ces corpus sont très divers, et les jeux de champs correspondant sont également très divers.

Néanmoins, il est nécessaire de comparer des documents d'un corpus à l'autre, donc de rapprocher certains de ces champs et de définir dans quelle mesure il est possible de les comparer. Comparer la datation d'un sceau et d'une statue a un sens, mais comparer leur poids n'en a probablement pas. Le rapprochement de ces champs est un point à prendre en compte de façon spécifique.

Là encore, le système doit donner la possibilité au chercheur de choisir librement les associations dont il a besoin.

Les corpus ne sont pas une structure suffisante pour regrouper les documents. Il a donc été décidé d'en créer une autre, plus polymorphe, qui permette de regrouper librement objets et documents en provenance de différents corpus.

Ces ensembles peuvent être créés et modifiés dynamiquement, et ils permettent aux chercheurs de rapprocher des documents ou objets en fonction d'une interrogation qui leur est propre. Cette structure a pour but d'être utilisée comme un outil de travail. Un chercheur enregistré a la possibilité de créer ses propres ensembles, accessibles de lui seul.

Les images, qui sont au cœur du projet, font l'objet d'une gestion particulière.

Tout d'abord, les images sont de natures, de formats et de qualités très diverses. Le but de MAVI est de réunir pour chaque document, autant d'images pertinentes que nécessaire, et de la meilleure qualité possible. Un outil logiciel a été développé spécifiquement pour permettre d'accéder en ligne à des images de très grande taille (supérieure à 10 millions de pixels) via des connexions faible débit.

Chaque image incorporée dans le système MAVI a un propriétaire identifié. Notre but est de garantir les droits de son propriétaire contre toute exploitation non autorisée des images. Cela peut aller de l'appropriation passive par un tiers, à la tentative de téléchargement massif d'un corpus entier. La première mesure que nous avons prise, est de n'autoriser l'accès aux images qu'au travers d'un système logiciel capable de prendre toutes les mesures de sécurité nécessaires □ authentication, autorisation, etc.

La question de l'archivage officiel de ces images, et de la protection des droits a été traitée □

- Chaque image est enregistrée auprès d'une autorité internationale et possède un numéro d'identification unique.
- Toute image envoyée vers un client est watermarquée. La signature contient entre autres informations, le nom du propriétaire de l'image et le numéro d'immatriculation de l'image. Ces informations sont répétées dans un en-tête écrit en clair.
- Les images sont visualisées dans un plug-in standard qui rend difficile l'enregistrement par un internaute indélicat.

MAVI n'a pas pour objet d'être propriétaire des images qu'il distribue, et de nombreuses conversations ont eu lieu avec les collaborateurs du projet sur ces points de sécurité.

L'exploitabilité d'une base documentaire repose sur la fiabilité scientifique des données qu'elle contient. Pour chaque corpus de données, MAVI s'est attaché la collaboration des spécialistes du domaine, il est donc important de garantir l'intégrité des documents qu'ils ont enregistrés dans la base.

De plus, la possibilité de procéder à ces enregistrements en ligne via l'Internet est rendue indispensable par le caractère international des collaborations établies.

Pour cela, un système de sécurité particulier sera mis en place, garantissant la sécurité et la fiabilité de ces données. Il reposera sur des échanges de clés et de mots de passe cryptés via des protocoles sécurisés, et par une vérification de l'identité de l'utilisateur avant chaque modification de la base.

Un système hiérarchique de droits d'accès sera mis en place. Un utilisateur ne pourra apporter de modification qu'à la partie de la base qui l'intéresse. Ce droit lui sera octroyé par un administrateur. Les modifications ne seront mises en ligne qu'après validation par une autorité scientifique d'administration, qui aura la charge de vérifier l'intégrité des données à cet instant-là.

Ce système sera indépendant des systèmes de sécurité mis en œuvre par le serveur de base de données que nous utilisons.

Toute modification de la base sera enregistrée par un système de surveillance, qui notera l'identité de l'indexeur et ce qui a été fait.

La question de la pérennité de ces données a également été soulevée dès les premières réunions de spécifications. Ce point est encore assez largement en question, tant il est difficile dans ce domaine de voir au-delà d'une dizaine d'années.

Afin de ne rien compromettre, nous avons fait le choix d'utiliser des formats de stockage de données les plus standards possibles (images JPEG, données exportables en XML). Ceci semble garantir pour une dizaine d'année l'exportation des données vers d'autres environnements logiciels que ceux utilisés aujourd'hui.

La pérennité du logiciel semble plus difficile à garantir sur un terme aussi long, elle est probablement moins importante, tant les concepts et les moyens dans ce domaine évoluent rapidement.

Ce point doit encore faire l'objet de réflexion et d'améliorations.

La construction d'une interface d'accès est un des points importants du projet, qui mérite d'être cité ici. La plupart des utilisateurs de ce système ne sont pas des spécialistes informaticiens, et leur connaissance de l'ordinateur se limite souvent à la connaissance d'outils de bureautique classiques.

Une réflexion importante est donc en cours sur la façon de concevoir une interface, tant pour l'indexation des données que pour leur consultation en ligne, qui soit à la fois

ergonomique, agréable à utiliser et rapide. Nous avons passé un accord de service avec une société extérieure, dont les compétences appliquées à des projets ressemblant au nôtre est reconnue.

#### Recherche de données

L'accès aux données pourra se faire suivant plusieurs modes□

- Un mode de requête classique, que de nombreux chercheurs manipulent déjà dans d'autres systèmes, et qu'ils souhaitent pouvoir utiliser sur MAVI.
- Un mode d'association automatique simple, qui est un ensemble de requêtes classiques préétablies, et accessibles sous forme d'hyperliens.
- Un moteur de recherche en langage naturel, produit commercial dont nous avons acquis une licence, et qui a fait ses preuves dans de nombreuses applications importantes. Une collaboration avec la société qui développe ce moteur est en cours de négociation, afin de l'adapter aux spécificités de MAVI.
- Un moteur de comparaisons d'objets utilisant les informations de la base, et fonctionnant en mode semi-automatique.

Ce dernier point fait encore l'objet de recherches. Nous comptons l'utiliser dans deux domaines□ celui de la découverte de hiérarchies naturelles dans la base, permettant la construction de thésaurus spécifiques, et celui d'associations entre objets dans des ensembles supplémentaires.

Ce type de recherche s'inscrit pour une certaine part dans le thème de «□indexation par le contenu□». Le L2TI s'intéresse à ce thème depuis plusieurs années déjà, et a montré l'originalité de son approche du problème. La façon dont le sujet sera abordé dépendra étroitement du type de données enregistré dans la base. Nous aurons à notre disposition de très nombreuses images, associés à des documents textuels précis et rédigés par des experts.

La recherche automatique de mots-clés sur des critères entropiques et dérivés sera la première chose que nous tenterons de mettre au point. La masse de documents textuels sera très importante, et disposer d'un outil de ce type permettra de tester la pertinence et le caractère discriminant de tel mot-clé, sur un corpus précis ou sur des ensembles plus larges.

Ce type de technique a déjà été mise en œuvre au L2TI sur des bases de tailles analogues, mais dans un domaine tout autre.

La mise en hiérarchie, réalisable par des approches multi-échelles, permettra de regrouper les ensembles de mots-clés sur plusieurs niveaux. Il sera intéressant de tester avec les spécialistes du domaine si ce type d'approche peut conduire à des outils d'aide à la construction de thésaurus spécifiques à la base, et le cas échéant à l'automatisation de la construction de ces thésaurus.

L'autre point sur lequel nous mènerons des travaux de recherche sera l'association automatique de documents textuels sur des critères de ressemblance établis en aveugle.

Si la comparaison de textes libres, même multilingues, est un problème dont les solutions commencent à émerger, celle de textes typés reste assez ouverte.

Nous allons disposer ici d'une base de taille importante, très structurée, et que les experts auront déjà largement structurée. Les *a priori* utilisables en classification et associations sont donc très forts, ce qui est assez rare dans ce domaine. Cette base constituera un champ d'investigation rare et qui sera très riche.

Il sera donc intéressant de prolonger les recherches menées au L2TI dans ce domaine, de façon à les appliquer à la problématique spécifique de cette base ☐

- Mise en association d'objets de types différents, avec correspondances de types.
- Construction de fonctions de ressemblances d'objets, de distances.
- Construction de graphes hiérarchiques d'objets.
- Regroupements d'objets au vu de ces graphes.
- Recherche automatique de prototypes au vu de ces regroupements.

Ces trois derniers points ont déjà été mis au point dans d'autres domaines, formellement ressemblant. Il sera intéressant d'étudier les prolongements possibles dans le cas des données MAVI.

### Recherche dans les images

La masse importante d'images dont nous disposerons semble aussi ouvrir des perspectives passionnantes dans ce domaine. Même s'il est un peu tôt pour décrire ces perspectives précisément, quelques grandes lignes se dessinent naturellement ☐

- Extraction automatique de motifs (bulles, sceaux, monnaies, tablettes).
- Association automatique de motifs.

- Recherche d'un motif dans une base, dans une image ou dans un dessin plus précis.
- Association de motifs à des données textuelles récurrentes et identification de formes.

Ces recherches s'inscrivent en droite ligne dans les schémas classiques d'analyse d'images □ extraction de primitives (les motifs), comparaison, classification, reconnaissance et identification de ces primitives.

## Partenariats

Un projet tel que le MAVI ne peut, par définition, aboutir que s'il est articulé avec des collaborations étroites avec les **Musées** du monde entier, qui acceptent de prêter leurs images et de transférer éventuellement leurs données (lorsqu'ils disposent d'une base de données). C'est très exactement le sens des contacts établis depuis plusieurs mois, et des accords qui ont d'ores et déjà été conclus □

France □ Musée du Louvre, Département des Antiquités orientales □ B. André-Salvini fait partie du Comité de pilotage □ un accord particulier a permis la prise de vue de nombreux petits objets (empreintes de sceaux sur argile) et un accord plus global est actuellement en discussion avec le Musée du Louvre.

BNF, Cabinet des Médailles □ photographies de monnaies royales et d'empreintes sur argile □ une campagne de photographie de sceaux a été réalisée dans le courant du mois de mai 2003, et un accord plus général sera conclu pour l'intégration de plusieurs milliers de monnaies, sur le modèle de l'accord passé avec le British Museum.

Angleterre □ La Direction du British Museum s'est montrée extrêmement favorable à la collaboration. Des accords ont été passés avec les Départements du Proche-Orient (John Curtis), d'Égyptologie (N. Spencer), de Grèce et des Monnaies (A. Meadows). Dans chacun de ces départements, un spécialiste est chargé de collecter images et données pour le MAVI et de collaborer avec José Paumard pour définir un processus d'exportation des données du BM vers la base de données du MAVI. Le processus est déjà largement avancé avec les Départements du Proche-Orient, d'Égypte et des monnaies.

Nous avons également un accord avec les National Museums of Scotland pour intégrer les images d'une stèle particulièrement remarquable trouvée à Saqqara (Égypte).

D'autres contacts sont menés par une collaboratrice anglaise du MAVI (une post-doc).

Pays-Bas □ Grâce à la collaboration d'un étudiant hollandais, Wouter Henkelmann (en thèse, membre du GDR 2538), nous avons passé un accord avec le Musée de l'Institut Néerlandais du Proche-Orient à Leiden □ une campagne photo spécifique a permis la collecte d'images haute-résolution d'empreintes sur argile □ un autre accord a été passé avec l'Allan Pierson Museum d'Amsterdam.

Allemagne □ un premier accord a été conclu avec l'Archäologische Museum de Munich, qui nous a fourni une couverture photo de remarquables peintures achéménides (encore inédites). L'un des collaborateurs du MAVI, Florian Knauss (GDR 2538), négocie actuellement avec d'autres musées allemands.

République Tchèque □ un accord de principe a été passé avec nos collègues de Prague, pour fournir photos et données de leurs fouilles de la tombe d'Udjahorresnet près de Memphis (Egypte).

Italie □ conversations en cours avec plusieurs musées. Des rencontres ont eu lieu en mai 2003, à l'occasion d'une conférence que P. Briant a donnée à Naples sur le projet MAVI.

Etats-Unis. Un accord a été passé avec l'University of Pennsylvania Museum (Philadelphie) pour les photos d'empreintes photographiées au début du siècle (conservées sur des plaques de verre). Un autre accord lie le MAVI et l'Oriental Institute de Chicago sur les photos des sceaux de Persépolis. Par ailleurs, un autre collaborateur du projet (Matt Stolper, Chicago) a obtenu des photos de deux musées de Yale, et une collaboratrice de Yale (Erica Ehrenberg) nous a fourni les photos d'une publication récente d'objets achéménides (nous avons, pour ce faire, également l'accord de l'éditeur allemand □ le Deutsche Institut d. Archäol.).

Iran □ un accord a été donné par la Direction du Musée archéologique de Téhéran.

Bien d'autres contacts ont été pris (par exemple avec le Miho Museum, Japon), et une réunion spécifique sera organisée au printemps 2004 pour faire un inventaire exhaustif des musées concernés, et pour désigner des correspondants dans chacun des pays considérés.

## **Conséquences attendues**

Les conséquences attendues sont de différents ordres □

- Comme on l'a indiqué dès le départ, à la suite du projet achemenet.com, le projet MAVI aura pour conséquence de créer un véritable réseau international de chercheurs dans un



domaine (l'histoire du Moyen-Orient unifié de l'Indus à la Méditerranée sous l'égide des Grands rois perses) qui n'est actuellement que fort peu représenté dans les Universités et les organismes de recherches (dont les axes «Antiquité» restent déterminés par les traditions gréco-romaines, égyptiennes et proche-orientales [avant les Perses], pour des raisons que P. Briant a souvent explicitées [voir *Leçon inaugurale au Collège de France*, Paris, 2000]). Le nombre et la qualité des chercheurs qui ont accepté de collaborer depuis l'année 2000 sont tout à fait impressionnants, de même que l'écho que le nouveau projet a suscité non seulement dans les Universités mais aussi dans les Musées du monde entier. C'est également dans cette vision que s'insèrent la création, par le collège de France, d'une chaire spécifique (1999) et la création, par le CNRS, d'une équipe internationale spécialisée (GDR 2538 «Réseau international d'études et de recherches achéménides»). En l'absence d'autres chaires de ce type dans quelque pays que ce soit ou d'autres équipes spécialisées, le réseau scientifique international doit se créer et fonctionner par l'intermédiaire de grands projets collaboratifs et par l'intermédiaire de la Toile.

- Du côté des recherches spécifiques sur le transfert des images, le Laboratoire L2TI attend des retombées extrêmement positives d'une telle collaboration. Ses activités de recherche s'articulent autour de plusieurs axes, dont la consultation de bases de données multimédia, qui intéresse MAVI. Ce projet est pour le L2TI une excellente occasion de tester, d'affiner et de faire progresser des concepts et des outils issus de travaux débutés il y a deux ans déjà. Ces outils se fondent sur des modélisations de type mélange de modèles afin de structurer les informations de façon originale. Ils ont été appliqués avec succès lors d'une collaboration interne à Paris 13, dans le but de classer des images de lignes de niveaux des eaux du delta interne du Niger. Cette base comportait environ 12 mille images, mais aucun texte associé. MAVI présente la nouveauté d'être une vraie base multimédia, de taille importante, dans laquelle les textes et les images sont indissociables, et pouvant être associée à des thésaurus. Les problèmes qui se posent sont donc nouveaux, tout en se plaçant dans la continuité de travaux déjà effectués au sein du laboratoire. Enfin, les résultats que nous obtiendrons pourront être confrontés à l'avis d'experts des données que nous allons traiter, ce qui est aussi un point extrêmement positif, et la garantie que nos recherches iront dans la bonne direction.

- Plus fondamentalement encore, le projet a pour perspective, très enthousiasmante, de créer une sorte de modèle de «Musée virtuel» associé à une base de données et à un moteur de recherches. Pour la première fois, tout le patrimoine culturel d'une grande civilisation sera rassemblé, archivé, organisé, consultable en ligne, grâce à la collaboration internationale et multidisciplinaire de chercheurs, d'institutions et de musées, que ce patrimoine ait été rassemblé par des dizaines de musées dans le monde, ou/et qu'il soit augmenté à la fois par les travaux d'analyse menés par des chercheurs, ou par des archéologues travaillant dans les pays considérés qui acceptent de mettre leurs découvertes immédiatement à la disposition des chercheurs par le biais du «Musée achéménide virtuel». Ce patrimoine unique ne sera plus dispersé entre de multiples musées, il ne sera pas invisible dans les réserves des musées, il ne sera pas non plus figé dans un état N, il sera au contraire réuni, accessible et en développement continu. Du point de vue scientifique, la collaboration attendue de dizaines de musées dans le monde est déjà une nouveauté, dans la mesure où ces musées accepteront aussi de fonctionner en dehors de la sphère marchande et s'intégreront plus étroitement encore au sein d'un projet de recherches international.

- Il convient enfin de souligner que le Musée Virtuel ne sera pas ouvert simplement aux chercheurs. Il sera accessible à tout le public intéressé. Une entrée spécifique sera créée, où le public intéressé mais non spécialiste pourra être introduit à l'espace-temps considéré, à la composition de ce patrimoine artistique et archéologique. Il pourra naviguer très librement dans le Musée, en suivant ses désirs propres qui ne sont pas nécessairement identiques à ceux du chercheur spécialisé. En bref, lieu de recherches, le MAVI a également vocation à devenir une institution culturelle ouverte et vivante.