



Colloque INTERNATIONAL

LANGUES d'EUROPE
ET DE LA
MÉDITERRANÉE

[HTTP://WWW.PORTAL-LEM.COM](http://www.portal-lem.com)

NICE

31 MARS – 2 AVRIL 2005

Fonctionnalités du portail LEM et choix technologiques

Jean-Pierre DALBÉRA*

Dans mon intervention, je présenterai la maquette du système d'information sur les langues que nous avons mis au point avec Henri Giordan, à partir de son rapport.

Ce dispositif a vocation à s'intégrer dans le futur portail d'information et de ressources du Musée des Civilisations de l'Europe et de la Méditerranée (MuCEM). Ce nouveau musée va être installé dans le fort Saint-Jean, un monument majeur du patrimoine historique bâti de Marseille et dans un bâtiment qui reste à construire, situé à proximité et dont l'architecte est Ruddy Ricciotti. Les réserves du musée seront situées dans un autre bâtiment qui sera construit près de la gare Saint Charles par l'architecte Corinne Vezzoni et associés.



Le fort Saint-Jean et l'entrée du Vieux Port de Marseille (photo MuCEM)

* Conseiller du directeur pour la recherche et la technologie. Exposé présenté avec Laurence Noël (MRT-MuCEM) et Gauthier Ubersfeld (Anyware).

La surface totale du musée sera de 25 000 mètres carrés. Le budget global du projet est de 150 millions d'euros. Le MuCEM ne sera pas un musée du passé, il sera vivant et traitera des sociétés contemporaines de l'Europe et de la Méditerranée, avec une démarche comparatiste. Il donnera une place importante aux expositions temporaires pour maintenir l'intérêt des visiteurs. Durant sa période de préfiguration qui a déjà commencé, de nombreuses expositions seront présentées. Ainsi, la prochaine manifestation sera consacrée cet été au hip-hop, un mouvement populaire qui s'est créé dans la rue il y a près de 30 ans.

Les collections du Musée National des Arts et Traditions Populaires et la collection d'Europe du Musée de l'Homme, soit près d'un million d'objets en tout, font actuellement l'objet d'un important chantier de récolement et de restauration, qui prépare le déménagement à Marseille.

Le musée possède également 500 000 photographies, des documents graphiques, plus de cent mille ouvrages, des milliers d'heures d'enregistrements sonores, les archives de toutes les recherches menées au sein du musée, plusieurs centaines de films. Ces documents ont été ramenés par les chercheurs lors de leurs campagnes ou ont fait l'objet de dons. À ces collections sur l'ethnologie de la France et de l'Europe, s'ajoutent celles appartenant aux pays du sud de la Méditerranée. Il s'agit soit de dépôts effectués par d'autres musées nationaux, soit d'acquisitions récentes du MuCEM.

Le centre de ressources du musée donnera accès à l'ensemble des documents originaux, mais privilégiera leur consultation sous forme numérisée. Pour le mettre en œuvre, il nous faut construire un système d'information, intégrant les bases de données et les catalogues numérisés des collections et des fonds du musée, et permettant la consultation de documents primaires numérisés. Il s'agit là d'un important chantier documentaire qui nécessite des moyens financiers que nous recherchons activement.

À l'origine, le MNATP a été conçu comme un musée-laboratoire, c'est-à-dire qu'un laboratoire de recherche en ethnologie de la France était associé au musée. L'organisation retenue pour le futur musée sera différente. Elle privilégiera un fonctionnement en réseau autour de trois pôles de recherche :

- Un pôle de recherche sur les sociétés de la zone euro-méditerranéenne et sur les collections d'objets associés,
- Un pôle consacré aux recherches et études sur la médiation et sur la muséographie/muséologie,
- Un pôle sur les applications culturelles des technologies de l'information et de la communication pour la documentation, la valorisation des recherches et l'édition multimédia.

Chaque pôle comprendra des laboratoires sous convention avec le musée et associera des universités et des organismes de recherche. L'Institut d'ethnologie méditerranéenne et comparative (IDEMEC), un laboratoire de la Maison méditerranéenne des sciences de l'Homme (MMSH), fait partie de nos partenaires privilégiés, mais des conventions sont en cours de signature avec plusieurs autres laboratoires dont l'UFR d'études slaves (université de Paris IV), le laboratoire « Modélisations et simulations pour l'architecture, l'urbanisme et le paysage » (CNRS, ministère de la Culture, Marseille), le laboratoire « Culture, Communication et Société » (ENS, Lyon), le LEDEN :

« Laboratoire d'Évaluation et de Développement pour l'Édition Numérique » (MSH Paris Nord, CNRS, Université Paris VIII).

Un Groupement d'intérêt scientifique (GIS), associant le CNRS, des laboratoires universitaires et plusieurs partenaires scientifiques et culturels est à l'étude.

C'est avec des chercheurs et doctorantes du LEDEN, dont fait partie Laurence Noël qui a travaillé sur le LEM, que nous avons lancé une politique multimédia au sein du MuCEM destinée à mieux faire connaître ses activités de recherche.

Une collection de publications multimédias en ligne vient d'être créée, avec le soutien de la Mission de la recherche et de la Technologie (MRT), sur les enquêtes ethnologiques et les collectes d'objets. Le premier numéro est consacré à « L'olivier, trésor de la Méditerranée »¹. Parallèlement, nous sommes en train de réaliser le site web de l'exposition « Hip Hop, art de rue, art de scène »² qui sera couplé à un espace interactif au sein de l'exposition et nous produisons avec le LEDEN le site de la dernière exposition des ATP, appelée « Lumières sur les ATP »³.

Les prochaines réalisations seront consacrées au travail du verre en Europe et en Méditerranée, au renouveau des musiques traditionnelles, aux cafés, etc.

J'en viens maintenant au projet sur les Langues d'Europe et de la Méditerranée : le LEM.

L'objectif est de réaliser un portail sur les langues, qui aura sa place dans le système d'information du musée. Le projet LEM s'inscrit naturellement dans le contexte du MuCEM puisque la langue est un élément majeur des cultures dont le musée possède les traces matérielles et immatérielles.

Le LEM est destiné au public, c'est pourquoi tous les outils développés doivent être simples d'utilisation, ergonomiques et accessibles. Le site doit être attractif, et les informations ne doivent pas être de nature encyclopédique. Il faut proposer aux internautes un moteur de recherche et un accès multilingue.

Les technologies utilisées pour la réalisation du portail doivent donc être choisies à bon escient et l'architecture du site doit être évolutive.

Le portail LEM doit intégrer un volet documentaire réalisé en collaboration avec les membres du réseau que nous avons l'ambition de constituer. En effet, Henri Giordan a proposé que le LEM soit un système capable de rassembler des informations provenant de différents partenaires, situés en France, mais également dans les pays d'Europe et de la Méditerranée. Parmi ces pays, l'Italie est un de nos partenaires privilégiés.

Au sein de ce portail, toute une machinerie doit fonctionner pour éditer les documents à distance, les indexer, les valider, les publier, les diffuser sur des supports variés, comme les mobiles de type PDA qui au moment de l'ouverture du musée, seront très répandus. Le portail LEM devra être interopérable avec le portail général du MuCEM pour faciliter les échanges.

¹ <http://www.musee-europemediterranee.org/olives/flash/index2.html>

² <http://hiphoponline.fr>

³ <http://www.musee-atp.fr/lumieres/>

Ce sont ces principes de base qui nous ont guidés pour bâtir la maquette du site web dynamique. On appelle dynamique un site dont les pages sont générées automatiquement et qui ne doivent pas être réalisées une à une lorsque des modifications ont été nécessaires.

Dans ce but, nous avons utilisé un système de gestion de contenus (CMS – Content Management System), qui permet, par exemple, à un spécialiste situé en Calabre d'écrire à distance sur le LEM, via internet, en utilisant son navigateur.

Un CMS permet de gérer des flux d'informations, des articles et des ressources (images, sons, cartes, références bibliographiques, extraits de films, fichiers PDF, etc.), de mettre des liens entre eux, d'attribuer des droits d'écriture, de valider des documents et des articles, de publier en ligne et de supprimer des informations, par exemple à une date donnée, d'éditer un bulletin électronique, d'envoyer des informations à des listes de correspondants, etc.

Des métadonnées (auteurs, langue utilisée, dates des différentes versions, commentaires, mots clés, etc.) peuvent être attachées aux documents afin de les retrouver ultérieurement avec le moteur de recherche.

Un système structuré, c'est-à-dire une base de données, a été incorporé à l'intérieur de la maquette pour traiter les informations de référence relatives aux langues, les adresses des centres de diffusion, des écoles, des librairies, des associations, les références bibliographiques. Ces éléments structurés sont indexés et peuvent être liés entre eux et aux articles.

Le système d'information réalisé offre également des possibilités d'import et d'export des données dans les formats XML ou CSV, ce qui facilite les échanges et l'actualisation des données, et garantit leur pérennité.

Le système dispose de fonctions d'impression et d'affichage en différents formats, tout particulièrement adaptés aux supports mobiles.

Une préoccupation des services scientifiques du ministère de la Culture est l'utilisation de logiciels sous licence GPL⁴.

Seul ce type de choix assure la pérennité du service public numérique, grâce à la maîtrise des sources des logiciels. Pour autant, il n'est pas si simple à mettre en œuvre dans le monde culturel car il nécessite l'existence d'une communauté stable de développeurs travaillant en réseau qui améliore les outils et les fait évoluer.

Pour évaluer les possibilités en ce domaine, Laurence Noël et moi-même, avons dans un premier temps étudié l'offre de CMS libres d'utilisation comme LENYA (écrit en Java) ou SPIP (écrit en Php).

Fondé sur Cocoon, LENYA⁵ aurait été utilisable à condition d'avoir des possibilités internes de développement informatique, ce qui n'est pas le cas au sein du musée. SPIP⁶ est plus facile d'accès mais n'offre pas en standard des

⁴ La Licence Publique Générale (General Public License) est destinée à garantir la liberté de partage et de modification des logiciels libres, et à assurer que ces logiciels soient effectivement accessibles à tout utilisateur.

⁵ <http://lenya.apache.org>

⁶ <http://www.spip.net>

fonctionnalités suffisantes pour satisfaire le cahier des charges du LEM et lui aussi impose des développements informatiques complémentaires.

Par contre, pour la partie base de données de la maquette, nous avons pu mettre en œuvre l'application XToGen⁷, développée à la demande de la MRT par la société Pass-technologies⁸ et qui est sous licence GPL.

Le montant des moyens mis à notre disposition par la DGLFLF et la Mission de la recherche et de la technologie (MRT) du ministère de la Culture a été un facteur limitant pour engager de nouveaux développements informatiques destinés à adapter un CMS libre et le coupler à la base de données, c'est ce qui nous a conduits à retenir, pour la maquette du LEM, une offre commerciale de CMS.

Le CMS utilisé a été développé par la société Anyware⁹, notamment à partir de Cocoon, l'application Xtogen y a été intégrée.

À ce stade, nous avons choisi de créer notre système en java-xml plutôt qu'en Php, ce qui nous a semblé un bon moyen de séparer la forme du fond, d'assurer l'interopérabilité à long terme et de gérer le multilinguisme et l'indexation. Toutefois, les outils libres en Php se développent rapidement, ils disposent d'une communauté de développeurs très importante et méritent également d'être testés dans des conditions d'utilisation réelle.

L'habillage graphique des pages du LEM et son adaptation au cahier des charges ont été effectués par Laurence Noël et un ingénieur d'Anyware, en collaboration avec Henri Giordan et moi-même.

C'est le résultat d'un travail collectif récent que nous vous présentons. La maquette demande à être testée avec des utilisateurs et évaluée pour réaliser une première version opérationnelle du portail.

La page d'accueil est composée :

- D'une fenêtre permettant le choix d'une langue ou d'un pays,
- D'une fenêtre d'interrogation d'un moteur de recherche,
- D'une série de rubriques cliquables : la boutique, la boîte à outils, le forum, les éléments d'enseignement, le réseau en lui-même,
- D'actualités sur les langues,
- D'un accès à des cartes présentant la répartition des langues en Europe.

La possibilité est offerte de choisir l'affichage en anglais ou en français, mais il pourrait se faire dans d'autres langues comme le permet la technologie XML-UNICODE. Ce n'est qu'une question de coût de réalisation.

Pour chercher une des langues proposées, on clique dans une liste déroulante. On arrive alors sur une page où l'on trouve, dans la partie supérieure, une barre de navigation avec les grandes rubriques de la page d'accueil et l'accès au

⁷ <http://xtogen.tech.fr>

⁸ <http://www.tech.fr>

⁹ <http://www.anyware-tech.com/fr/index.html>

moteur de recherche. Sur la gauche, la rubrique « identification » est destinée aux rédacteurs et administrateurs du site.

Quand une langue est choisie, l'alsacien par exemple, on arrive sur la fiche de la base de données correspondante. On y découvre un certain nombre de champs, dont on constate, que certains sont interactifs. Il s'agit de champs indexés par le système. Ainsi, on peut par exemple accéder à toutes les fiches des langues traitées dans la base appartenant à une même famille linguistique.

On peut visualiser des documents visuels, des cartes, ou écouter des enregistrements sonores. Au bas de cette fiche, se trouve un lien appelé « En savoir plus ». Henri Giordan a souhaité offrir au visiteur la possibilité d'accéder à des textes beaucoup plus développés, comportant éventuellement des tableaux. C'est dans ce cadre qu'il est fait appel au CMS, sans pour autant que l'internaute s'en rende compte. On accède alors à un ensemble de documents multimédias dont le sommaire figure en marge de gauche.

Quelle que soit la page que l'on consulte, le bas de l'écran présente toujours le chemin que l'on a emprunté depuis la page d'accueil pour y arriver. Même si tout ceci est apparemment semblable à ce qu'on trouve sur le web, il convient de souligner qu'il ne s'agit pas d'un site statique, mais d'un site dynamique intégré, qui dispose de fonctionnalités très puissantes.

Lorsqu'on accède par exemple à la rubrique « Actualités », on peut visualiser des documents multimédias et imprimer la page au format PDF. Cette solution permet d'obtenir une page imprimée bien présentée. Pour la consultation des cartes, une fonction de téléchargement permet l'agrandissement, ou l'impression. La maquette du LEM fonctionne, elle reste à tester, pour l'améliorer en fonction des besoins réels des utilisateurs.

Nous allons maintenant présenter la partie du LEM réservée à la saisie en ligne des informations avec gestion des flux (rédaction, validation et publication). C'est ce processus qui est l'origine du choix technologique d'un CMS afin de respecter le caractère fédératif du réseau conçu par Henri Giordan.

Après avoir tapé les codes d'accès, on arrive à un écran réservé à la saisie des données. L'interface est un peu plus complexe, mais elle est destinée à un nombre réduit d'utilisateurs qui peuvent recevoir une petite formation.

La base de données, réalisée sous l'applicatif Xtogen comprend trois ensembles de données qui peuvent être liés entre eux : un ensemble de fiches pour les descriptions des langues, un pour les adresses, un pour les références bibliographiques.

En sélectionnant le mode saisie pour les fiches sur les langues, on accède à un écran dont plusieurs champs possèdent des listes déroulantes. Le système retenu permet la création et la modification de listes déroulantes en fonction des besoins, ce qui facilite la saisie en limitant le risque d'erreur. Dans le champ « document sonore », on peut attacher aux données un fichier son, aussi bien qu'un fichier image. Tous ces champs sont aisés à remplir. Les données saisies doivent être ensuite validées par le comité scientifique dont Henri Giordan vous a déjà parlé.

Pour la saisie d'informations non structurées comme des notes ou des articles complets, on utilise les outils du CMS. Les documents sont saisis en ligne avec un éditeur de traitement de texte standard, qui prend en charge toutes les typographies, arabe, grecque, etc. L'enregistrement des modifications permet la

saisie d'un commentaire sur le document principal, comme par exemple la justification d'une modification apportée.

Le CMS mémorise la date de rédaction et gère les versions successives d'un document. Cette fonctionnalité est importante, car le travail en réseau nécessite l'enregistrement de la chronologie des échanges entre les équipes travaillant sur les langues et les différents comités scientifiques. La modification d'un document est vérifiée par le comité puis validée par le titulaire des droits d'administration qui peut alors le mettre en ligne. L'actualisation du site web est ainsi beaucoup plus rapide que pour des sites statiques, car les pages affichées sont générées et modifiées automatiquement. C'est là que réside la différence avec la méthode classique de création de pages web.

Ce système professionnel est d'ores et déjà utilisé par les grands groupes industriels, autant pour leur communication externe qu'interne. À partir des informations contenues dans le CMS, on peut extraire les documents au format PDF, pour les imprimer ou les envoyer sur un autre site web, sur un téléphone portable, ou sur un PDA.

Cette maquette constitue un premier pas dans l'élaboration de l'outil d'informations et d'échanges sur les langues d'Europe et de la Méditerranée, en conformité avec les grandes orientations définies par Henri Giordan.

Faute de moyens financiers suffisants et de personnels permanents affectés au projet au sein du musée, l'architecture informatique de la plate-forme n'est pas encore stabilisée et de nombreuses fonctionnalités restent à mettre au point et à tester, comme par exemple la gestion des références bibliographiques. Des études complémentaires sont nécessaires pour l'adapter aux besoins d'utilisateurs réels et en faire un outil opérationnel au service d'un projet international. Malgré ses insuffisances, cette maquette permet de mieux faire comprendre les orientations du LEM et d'engager une recherche de nouveaux partenaires avec davantage d'efficacité