

## **Stage de Recherche 2008**

### **Master *Recherche en Informatique*, deuxième année**

## **Université Montpellier II**

### ***Laboratoire***

- Nom du laboratoire LGI2P-EMA
- domaine d'activité : Génie Informatique et Ingénierie de Production
- adresse : Parc Scientifique Georges Besse  
F - 30035 Nîmes Cedex 1
- telephone : 04 66 38 70 00
- fax
- 
- URL : <http://www.ema.fr>

### ***Tuteur laboratoire***

- Nom : Michel Crampes (HDR Montpellier II)
- laboratoire/équipe: LGI2P-EMA/ équipe KID
- thématique : web sémantique, cartographie des connaissances, ontologies, indexation multimédia, composition multimédia automatique, modèles bayésiens
- e-mail : [michel.crampes @ema.fr](mailto:michel.crampes@ema.fr)
- URL : <http://www.ema.fr/~mcrampes/>

## Sujet de Stage : scénarisation dynamique d'agents

- mots-clés :

connaissances, architecture multi-agents, web sémantique ontologie, cartographie, info-vis, multimédia.

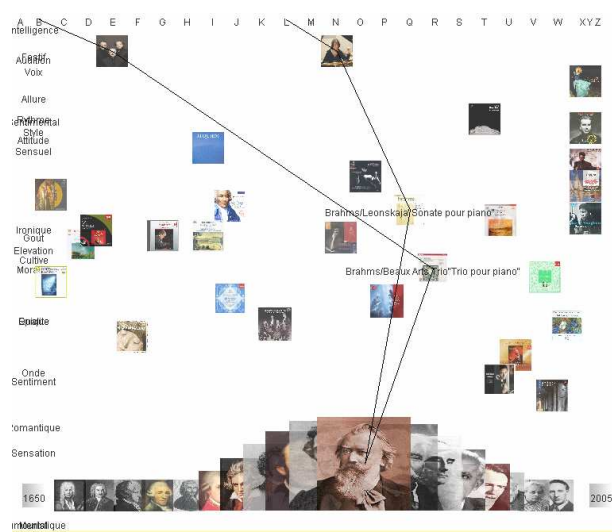
- modules du Master requis (consulter la page web)

(préférentiellement, mais pas obligatoire) Intégration de Données XML, Représentation des connaissances, modèles multi-agents, Visualisation d'Informations, Apprentissage

- Contexte

Les nouvelles technologies de création d'objets dématérialisés (photos, vidéos, documents numériques), et leur partage en ligne ont transformé l'enjeu de la rareté des objets matériels (par exemple l'unicité de la photo argentique), en un enjeu de la rareté sensorielle et cognitive, c'est-à-dire de la capacité dont peut disposer un individu ou un collectif d'individus à appréhender l'offre massive à sa disposition. Ceci vaut aussi bien pour les documents en ligne que pour la connaissance partagée dans un domaine donné.

Deux approches majeures peuvent être identifiées pour faire face à ce nouvel enjeu : les moteurs de recherche du type Google qui répondent par une liste d'objets à une requête non structurée, et la cartographie qui permet de naviguer sous forme assistée dans une représentation visuelle de la masse d'objets pour trouver ceux qui répondent le mieux au besoin.



Les deux approches ont un problème commun : l'indexation des objets soit pour les retrouver pour la première, soit les organiser pour la seconde. Ceci constitue un verrou, surtout dans le domaine de l'image ou de la vidéo. Par contre elles se différencient dans leur stratégie. Les moteurs fouillent la base pour proposer des listes, souvent elles-mêmes massives, éventuellement ordonnées, mais sans, ou avec peu de traitement sémantique. La recherche d'information se fait donc en aveugle, sur la seule foi des résultats des moteurs de recherche. De plus, d'une requête à l'autre, il n'y a pas de trace de l'évolution sémantique des résultats.

Le grand intérêt de la cartographie est de fournir à l'utilisateur une vue globale du domaine, et ainsi

de situer les réponses aux requêtes dans leur contexte. On peut rechercher des voisinages, et mesurer la distance de certains objets à la requête. Dans ce but, les environnements de cartographie doivent disposer de stratégies d'organisation des objets basée sur la sémantique et proposer des assistants à la navigation. Mais les différents environnements de cartographie actuellement proposés sont relativement statiques et faiblement pilotés par la sémantique. Les applications réellement opérationnelles sont rares. Ils répondent peu au besoin alors qu'ils devraient apporter de nouvelles solutions comparées aux solutions proposées par les moteurs de recherche.

- Position du problème

Depuis plusieurs années nous avons développé un environnement de cartographie dynamique basé sur trois technologies : des techniques issues des communautés 'graph drawing' et 'infovis', un modèle multi-agents, et une formalisation qui utilise les cadres formels du Web sémantique (XML, XML schema, XSLT, OWL, RDF, JAXB). Cela nous permet de disposer de stratégies de projection à deux dimensions de mondes parfois très riches. Mais la flexibilité de l'environnement a une contrepartie : il devient impossible à un non expert d'élaborer des

stratégies efficaces de projection visuelle, et de mettre en œuvre l'environnement étant donné sa complexité tant conceptuelle qu'ergonomique.

- Ces difficultés étaient attendues. La réponse était déjà connue aussi. Elle consiste à faire cohabiter trois principes : 1) nous avons développé un moteur de scénarisation d'agents avec un langage de scénario associé : les objets à représenter et à mettre en espace reçoivent des instructions synchronisées décrites dans un langage 'agent' dit de 'scénario', et basé sur un cadre sémantique formel (RDF et schéma XML) ; la philosophie 'agents' est donc très marquée. 2) un scénario de mise en espace peut être décomposé en sous-scénarios (motifs) qu'il est possible d'assembler différemment pour disposer d'autres stratégies de mise en espace. 3) Il est possible de visualiser ces motifs et leur assemblage sous forme de graphes eux-mêmes manipulables de manière récursive par le même principe de mise en scène.
- Travail demandé
- Le principe du moteur et de son langage sont acquis et implémentés. Plusieurs scénarios ont été réalisés et même pour l'un d'entre eux breveté.
- Les principes 2 (assemblage de motifs) et 3 (visualisation et auto représentation) sont à développer en tenant compte du fait qu'il s'agit d'une architecture agents et que des aspects de communication inter-agents et de synchronisation sont à prévoir. Le langage, bien formalisé par un schéma, sera sans doute étendu pour mieux tenir compte de cet aspect.
- Il existe aussi un quatrième principe qui reste à mieux préciser : Supposons que l'on veuille mettre en scène une collection d'albums virtuels de musique avec leur description sémantique. Cet univers peut être décrit par rapport à un modèle formel UML ; il serait bon de générer automatiquement des mises en espace les plus efficaces en rapport avec l'expérience accumulée pour d'autres collections (photos par exemple). Ce procédé est décrit dans l'article [8]. Une forme d'apprentissage est-elle requise, ou bien suffit-il de considérer des formes de règles de transformation XSLT paramétrée ? Des essais dans cette seconde direction ont été menés avec succès, mais la première solution n'est pas à exclure et est à étudier
- Possibilité de poursuivre en thèse

Possibilité de bourse de thèse à l'EMA selon les résultats

#### Lieu et contexte de travail

Le stage pourra se faire dans les locaux du LGI2P de l'Ecole des Mines d'Alès à Nîmes. Le stagiaire sera en relation étroite avec un étudiant en stage de master I de l'Université UNSW de Sydney déjà identifié (avec dialogues fréquents en Anglais).

## Références bibliographiques

- [1] **Using Concept Lattices for Visual Navigation Assistance in Large Databases: Application to a Patent Database** (extended version). Villerd, S. Ranwez, M. Crampes, D. Carteret. Journal of General Systems, special issue on Concept Lattices and Applications (accepted)
- [2] **Using Concept Lattices as a Visual Assistance for Attribute Selection**. J. Villerd, S. Ranwez, M. Crampes. 16th International Conference on Conceptual Structures, ICCS'08, Toulouse, France, July 7-11 2008.
- [3] **Cartographie sémantique auto-organisée d'un référentiel de connaissances partagé**. M. Crampes, S. Ranwez, J. Villerd. 19èmes journées francophones d'Ingénierie des connaissances, IC'08, Nancy, France, June 18- 0 2008.
- [4] **Navigation sur des Cartes de Connaissances supportées par un Treillis de Galois**. Colloque. J. Villerd, S. Ranwez, M. Crampes. Carto 2.0, Noisy-le-Grand, France, April 3 2008
- [5] **Using Concept Lattices for Visual Navigation Assistance in Large Databases: Application to a Patent Database**. Jean Villerd, Sylvie Ranwez, Michel Crampes, David Carteret, Jean-Michel Penalva. In [CLA 2007](#), the fifth International Conference on Concept Lattices and Their Applications. Montpellier, France, October 24-26, 2007.

- [6] **Cross-linking Music and Pictures through Moods.** Michel Crampes, Olivier Gout, Nicolas Mille, Jean Villerd, Sylvie Ranwez. In proceedings of Multimedia Systems and Applications, [MSA'07](#), within the scope of the 2007 World Congress in Computer Science Computer Engineering, and Applied Computing, [WORLDCOMP'07](#), Monte Carlo Resort, Las Vegas, Nevada, USA, June 25-28, 2007.
- [7] **Automatic Playlist Composition in a Dynamic Music Landscape.** Michel Crampes, Jean Villerd, Andrew Emery, Sylvie Ranwez. International Workshop On Semantically Aware Document Processing And Indexing, [SADPI'07](#), In Cooperation with ACM SIGWEB, Montpellier, France, 21-22 mai 2007.
- [8] **Concept Maps for Designing Knowledge Maps.** *Crampes, M., Ranwez, S., Emery, A., Velickovski, F. Mooney, C., Mille, N.* In Journal of Information Visualisation, vol 5, n°4(2006).
- [9] **i2dee : An Integrated and Interactive Data Exploration Environment Used for Ontology Design.** Fabien Jalabert, Sylvie Ranwez, Vincent Derozier and Michel Crampes. In [Managing Knowledge in a World of Networks](#), Lecture Notes in Computer Science, publisher: Springer Berlin / Heidelberg, Volume 4248/2006, ISBN 978-3-540-46363-4, pp. 256-271, 2006.
- [10] **Ontological ISA-Distance Measure for Information Visualisation on Conceptual Maps.** Sylvie Ranwez, Vincent Ranwez, Jean Villerd, Michel Crampes. In [On the Move to Meaningful Internet Systems 2006: OTM 2006 Workshops](#), Lecture Notes in Computer Science, publisher: Springer Berlin / Heidelberg, Volume 4278/2006, ISBN 978-3-540-48273-4, pp.1050-1061, 2006.
- [11] **Cartes conceptuelles pour l'ingénierie de cartes de connaissances adaptatives.** Crampes, M., Villerd, J., Velickovski, F., Mille, N. et Ranwez, S. Conférence Ingénierie des Connaissances, Nantes (2006).