

Temis lance Luxid, moteur de recherche du 'text mining'

« Solution innovante... véritable rupture technologique... création de valeur... » Les fondateurs de la « jeune pousse » Temis soignent leur marketing. Et cela s'avère efficace!

Créée en septembre 2000 à Paris et à Heidelberg, par quatre anciens cadres d'IBM Europe et un chercheur universitaire allemand, tous spécialistes du « *data mining* », Temis semble bien tenir ses promesses, toutes ses promesses.

Après ses deux années de décollage et de coopération avec Xerox, son organisation a été renforcée au sommet par l'arrivée d'Eric Brégand (ex XRT et Business Objects) et de Gilles Pouzenc (fonds d'investissement I2P: Wincorp, CBA...).

Aujourd'hui les références « grands comptes » s'accumulent et le chiffre d'affaires suit [lire ci-après].

Le '*data mining*'? « C'est l'une des techniques qui alimente la 'Business Intelligence' (système décisionnel); il traite de grands volumes de données numériques, utilise différentes méthodes issues des mathématiques, particulièrement des statistiques. Il automatise la recherche de relations, de corrélations, de dépendances et d'associations entre les données (...) »

Et le « *text mining* » ? Suite logique du précédent « *il transforme tout texte libre (courriers de réclamations, questions ouvertes dans un questionnaire, articles...) en une série de données analysables* ». Il repose sur « *l'analyse morpho-syntaxique (catégories grammaticales, identification de groupes nominaux...)* et sur *l'analyse sémantique (synonyme, recherche d'expressions, extraction de connaissance) (...)* ».

Luxid, solution d'analyse de l'information, est constituée de 3 couches applicatives:

- Luxid annotation factory: une palette d' « extracteurs » d'informations multilingues, identifiant les entités et leurs relations; il existe XeLDA, Insight discoverer extractor (reconnaissance de 50 formats ou 'corpus', analyse morpho-syntaxique et extraction de connaissance ou « règles d'extraction »...
- Luxid information mart: la plate-forme qui fédère des sources d'informations hétérogènes, enrichit les documents pour créer des bases de connaissance
- Luxid information analytics: une application Web dédiée à la découverte et au tri de l'information. Ce module trie, filtre en naviguant sur Internet; il fournit des analyses temporelles (il classe par rapport aux dates) et établit des analyses croisées. En synthèse, il peut fédérer en grappes ('clusters') les documents apparentés.

Luxid est une solution modulaire (ces trois sous-ensembles peuvent être acquis séparément); elle repose sur l'architecture UIMA d'IBM (*Unstructured information management architecture*).

L'investissement pour la version « grand compte » est de l'ordre de 100 K-euros. Il existe des bibliothèques prêtes à l'emploi ou « cartouches (« *skill cartridges* »), à 50 à 75 K-euros/an, soit généralistes, soit spécialisées (Sciences de la vie, biotechnologie, médecine / médicaments...), avec

option multilingues (jusqu'à 20 langues, dont le japonais, le chinois, l'arabe). Un kit de développement 'Studio' permet de développer ses propres « cartouches », avec son propre dictionnaire.

Temis travaille également à une version « *desktop* » (poste client seul) de son offre, pour des licences de 5 à 10 K-euros.