



Solutions in language processing

Lingway : 2 projets labellisés Technolangue

Paris, le 16 juin 2003 - Lingway, société spécialisée dans la conception et le développement d'applications basées sur les technologies linguistiques, annonce que 2 de ses projets viennent d'être labellisés dans le cadre du Programme Technolangue qui regroupe à l'initiative du Ministère de la Recherche et des Nouvelles Technologies, plusieurs acteurs de l'ingénierie linguistique en France, tant universitaires qu'industriels, et des utilisateurs (Bibliothèque Nationale de France, AFNOR, ...) autour du développement d'outils de traitement de la langue.

Ces deux projets baptisés WATSON (Web : Analyse de Textes, Sélection et Outils Nouveaux) et TILT (Trésor Informatisé de la Langue Technique) font partie du groupe de projets AGILE coordonné par Bernard Normier, Président de Lingway.

L'objectif du projet WATSON est de développer, adapter, intégrer des outils logiciels linguistiques notamment de structuration logique de pages Web, des « taggeurs », des « chunkeurs », des modules de reconnaissance d'entités nommées, de marquage textuel, des extracteurs, des « wrappers », des catégoriseurs, des résolveurs de co-référence et des « résumeurs », pour permettre le traitement de très grands corpus. Ces outils seront évalués dans le cadre de l'application expérimentale pour le dépôt légal du Web français, avec des corpus de tailles extrêmement larges et de typologies très variées. Ce projet permettra à Lingway de renforcer son offre de « Webmining ».

Les partenaires de WATSON sont : **La Bibliothèque Nationale de France**, www.bnf.fr, **le CNRS Laboratoire LATTICE-Talana**, www.talana.linguist.jussieu.fr, **le CNRS Laboratoire LaLICC**, www.lalic.paris4.sorbonne.fr

L'objectif du projet TILT est de constituer un corpus du français technique utilisé dans les normes et d'aligner ce corpus sur son équivalent anglais. Une brique logicielle linguistique de base sera constituée et permettra le développement d'applications linguistiques à forte valeur ajoutée, comme l'assistance à la traduction, la recherche d'informations complexes, l'indexation automatique ou la réalisation de résumés automatiques. Le corpus issu du projet sera ensuite étendu au 30 000 normes européennes actuellement disponibles.

Les partenaires de TILT sont : **L'AFNOR**, www.afnor.fr, **l'ATILF**, www.atilf.fr

A propos de Technolangue : www.technolangue.net

Technolangue est une action interministérielle dans le domaine du traitement informatisé des langues (TIL), lancée en 2003 et pilotée par le ministère chargé de la Recherche et des Nouvelles Technologies (Direction de la Technologie). 27 projets ont été labellisés sur 52 reçus.

Ce domaine, également connu sous le nom de « traitement automatique des langues » (TAL) ou « ingénierie linguistique », vise à fournir les technologies de la langue écrite et orale qui sont au cœur des applications de traitement de l'information et de la communication.

L'objectif principal de Technolangue est de mettre en place de manière pérenne une infrastructure de production et diffusion de ressources linguistiques et logicielles, d'évaluation des technologies de la langue écrite et orale, de participation aux instances nationales et internationales de normalisation et de standardisation et de veille informationnelle sur le domaine.

Technolangue est une action commune aux trois réseaux de recherche et d'innovation technologique (RNRT – Réseau National de Recherche en Télécommunications, RNTL – Réseau National des Technologies Logicielles, RIAM – Réseau pour la Recherche et l'Innovation en Audiovisuel et Multimedia), financée dans un cadre interministériel.

Les ministères concernés sont le Ministère chargé de la Recherche et des Nouvelles Technologies, le Ministère délégué à l'Industrie et le Ministère de la Culture et de la Communication.