

LE LOGICIEL DE GESTION TERMINOLOGIQUE GesTE

Dorina CHIȘ, Valeriu Dorin CHIȘ
 Tibiscus University, Timișoara
 Politehnica University of Timișoara

1. L'appel de l'Union Latine pour la création d'un software de gestion terminologique

Union Latine a lancé, le 24 juin 2006, un appel pour la création d'un software de gestion terminologique gratuit et conçu sur le modèle du code en source libre.

On a tenu ainsi à mettre à la disposition de tous les spécialistes oeuvrant dans le domaine de la terminologie un outil basé sur les normes internationales (ISO) et destiné à la gestion des travaux terminologiques. Le logiciel devait combler notamment un manque d'instruments de ce genre pour le traitement du roumain, et, par conséquent, l'application devait être conçue en roumain. Il fallait, toute fois, en prévoir des fonctionnalités multilingues, dans la perspective d'une traduction/localisation ultérieure.

Nous avons répondu à cet appel et nous avons réalisé le gestionnaire de terminologie GesTE.

2. Approche

Dans une première étape nous avons consulté des travaux théoriques du domaine (dont nous avons énuméré quelques uns dans la liste de la bibliographie sélective), ainsi que les bases de données terminologiques en accès libre, telles que :

- Banque de données terminologiques - Service de la langue française, de la Communauté française de Belgique, <http://www.cfwb.be/franca/bd/bd.htm>,
- Base de terminologie du CILF (Conseil international de la langue française) <http://www.cilf.org/bt.fr.html>,
- CRITER (Corpus du Réseau Interministériel de Terminologie), La délégation générale à la langue française et aux langues de France <http://www.culture.gouv.fr/culture/dqif/terminologie/base-donnees.html>,
- EURODICAUTOM – actuellement IATE (Inter Active Terminology for Europe) <http://iate.europa.eu/iatediff/SearchByQueryLoad.do?method=load>,
- ILOTERM – Base de données de l'Organisation internationale du travail, <http://www.ilo.org/iloterm/>
- IMF Terminology, Base de données terminologiques du Fond Monétaire International, http://www.imf.org/external/np/term/index.asp?index=fr&index_langid=2,
- Le Grand dictionnaire terminologique de l'Office québécois de la langue française, http://www.granddictionnaire.com/btml/fra/fr_motclef/index1024_1.asp,
- TermRom - CIMEC, România, <http://www.cimec.ro/PaginiGazduite/TR/default.htm>
- UNESCOTERM, Base de données terminologiques de l'UNESCO, <http://termweb.unesco.org/>
- UNTERM, United Nations Multilingual Terminology Database, <http://157.150.197.21/dqaacs/unterm.nsf..>

On a pu ainsi prendre connaissance de la manière dans laquelle les informations sont organisées et mises à la disposition des utilisateurs.

D'autre part, le bureau de Bucarest de l'Union Latine a organisé plusieurs rencontres consacrées aux problèmes de structure des bases terminologiques et à la conception des fiches. Ces rencontres nous ont donné l'occasion de constater que, selon le public destinataire et les instruments informatiques utilisés, les structures proposées par les participants variaient, tout en gardant un certain nombre d'éléments communs. Cela nous a fait comprendre qu'il fallait concevoir notre gestionnaire comme un instrument très flexible, qui puisse offrir à l'utilisateur non seulement une structure conçue conformément aux normes internationales, mais aussi la possibilité de la modifier, de la compléter et même de créer sa propre structure, peut-être mieux appropriée au public cible choisi et aux buts envisagés.

3. Caractéristiques du gestionnaire de terminologie GesTE

3.1. Support informatique

L'application utilise des programmes gratuits et compatibles avec les principaux systèmes d'exploitation acceptés sur le marché roumain. De plus, elle est conçue sur le modèle du code source libre, les utilisateurs étant libres de la modifier et de la distribuer, à la différence des logiciels développés par les sociétés commerciales.

GesTe peut être consulté en ligne, en Intranet ou installé sur son ordinateur personnel.

Les opérations sur la partie serveur sont effectuées par un script PHP qui interroge une base de données MySQL. Celle-ci comprend aussi bien la structure de la base (l'organisation des champs et des groupes, des listes et des entrées), les données terminologiques proprement dites et les informations concernant les utilisateurs et leurs possibilités d'accès (restreint ou total) à l'application.

La table système des données terminologiques n'a pas de structure fixe, étant modifiée par le programme en fonction des structures qui sont créées durant le travail. Pourtant, chaque champ de toute structure créée doit être associé avec une colonne de cette table.

Les interrogations MySQL sont réalisées côté serveur par un script PHP. Les informations sont ensuite transmises au client en format XHTML. Toutefois, on a séparé le code XHTML du code PHP, ce qui assure à l'application une grande flexibilité de fonctionnement, ainsi que la possibilité de procéder à des modifications à partir du niveau de l'interface utilisateur. Le formatage des données est réalisé à l'aide de fichiers .tlp qu'on peut éditer pour afficher les données en utilisant n'importe quel format sans modifier les algorithmes du programme.

La navigation d'une page à l'autre se fait par les paramètres GET, envoyés au fichier index.php. C'est ainsi que les accès vers d'autres fichiers du système restent cachés. De même, les paramètres utilisés pour les commandes ne peuvent pas dévoiler d'aucune manière la structure de la base de données et ne permettent pas d'attaque du genre *MySQL Injection*.

Au moment où le compte d'un nouvel utilisateur est créé, ses données sont enregistrées dans une table, accompagnées d'un identificateur. Par la suite, toute référence à l'utilisateur se fait par l'intermédiaire de l'identificateur créé au moment de l'enregistrement. Le mot de passe utilisé pour l'authentification n'est pas enregistré dans la base de données (et on ne peut le récupérer). En échange, le programme enregistre l'identificateur (hash md5) du mot de passe et le compare à

ceux des mots de passe introduits au cours des authentifications antérieures. De cette manière, un accès à la base de données MySQL, qui n'emploie pas l'interface de l'application, ne dévoilera pas les données d'identification des utilisateurs.

L'application permet l'exportation et l'importation des données terminologiques. Pour l'exportation un fichier en format propre (.gte) est généré. Celui-ci contient le résultat d'une recherche dans la base. Pour l'importation on utilisera le fichier (.gte) antérieurement exporté, ainsi qu'une structure compatible avec les données importées. L'exportation et l'importation permettent la sélection des champs des entrées sur lesquelles on va opérer. C'est ainsi qu'on peut exporter/importer seulement une liste de termes, seulement les termes et leurs définitions ou bien d'autres catégories de données.

3.2. Normes utilisées

Les normes utilisées dans l'élaboration de la base de données GesTE sont :

- ISO 1087 – concepts utilisés en terminologie ;
- ISO 12618 – principes de création des bases de données terminologiques ;
- ISO 12620 – catégories de données en terminologie ;
- ISO 12200 ou MARTIF – *Machine-Readable Terminology Interchange Format* – principes de conversion et programmation XML des banques de données terminologiques.

3.3. Langues de travail

L'interface et la structure de la base de données sont disponibles en catalan, espagnol, français, italien, portugais et roumain. La traduction des textes a été assurée par le bureau de Paris de la Direction de Terminologie et Ingénierie Linguistique de l'Union Latine.

L'utilisateur peut choisir la langue qui lui convient pour l'interface de travail et il recevra dans cette langue toutes les informations relatives à un certain terme (définition, contexte, explications...), même si le terme en cause appartient à une autre langue.

3.4. Structure de la base de données

La structure implicite de la base de données inclut les catégories d'informations indiquées dans les normes ISO 12620. Le champ *langue à laquelle appartient le terme* y a été ajouté, justement pour pouvoir offrir la possibilité évoquée plus haut, c'est-à-dire de choisir une autre langue de travail que celle à laquelle appartient le/les terme(s) en question.

Les noms et les descriptions des champs, chaque mot/fragment de texte, tout peut être édité et ensuite reformulé, modifié par ajout ou rejet d'éléments. Cela est valable pour les traductions aussi.

Chaque champ dispose d'une description claire qui exclut toute confusion.

La fiche terminologique est générée conformément à la structure active au moment du travail. (S'il y a plusieurs structures, il faut marquer et ensuite sélectionner la structure active.)

Pour la liste des domaines auxquels les termes appartiennent nous avons utilisé le modèle d'une ontologie employé dans la constitution de la banque terminologique d'Italie, modèle développé par Ingetraut Dahlberg.¹

3.5. Fonctionnalités

GesTE offre la possibilité de collecter, enregistrer, traiter, importer et extraire les données. On peut effectuer des recherches simples ou avancées (au moyen de filtres), ajouter ou effacer des champs, créer, consulter, valider ou supprimer des fiches, éditer les textes (définitions, exemples, contextes, explications), les modifier, les traduire, administrer la base. La navigation peut se réaliser par recherche de terme (simple, avancée) ou bien d'une fiche à l'autre, par des renvois.

Le logiciel est accompagné d'un **Guide d'emploi**, offert dans toutes les langues de travail, qui met à la disposition des utilisateurs toutes les informations nécessaires.

4. Conclusions

Le gestionnaire de terminologie GesTE met à la disposition des intéressés un instrument flexible et efficient, à l'aide duquel on peut collecter, traiter, stocker, exporter, importer les données terminologiques. Chaque utilisateur a la possibilité

d'adapter cet instrument à ses propres besoins depuis l'interface de travail, sans avoir à consulter un spécialiste en informatique.

La réalisation de ce logiciel est due à l'initiative et au support de l'Union Latine, dont le bureau de Paris l'utilise d'ailleurs avec succès, depuis l'année passée (2007). Il est présenté sur la page web de la Direction pour Terminologie et Ingénierie Linguistique de l'Union Latine, <http://dtil.unilat.org/ro/index.htm>.

L'utilisation de GesTE permettra la création et la diffusion de glossaires informatisés et devrait faciliter l'harmonisation des données terminologiques.

Notes

¹ Tel qu'il est présenté dans l'article de Giliola Negrini, « Quels systèmes conceptuels peut-on préconiser pour les banques de données terminologiques en ligne? », publié dans la revue *Terminologies nouvelles*, nr.18, juin 1998, *Terminotique et documentation*, pp. 41-53.

Références bibliographiques

1. Aussenac-Gilles, Nathalie, « Construction de modèles conceptuels à partir de textes », <http://www.infres.enst.fr/rdc05/aussenac.pdf> ;
2. Blampain (D.), Descotte (S.), Husson (J.-L.), Rohde (H.), Romary (L.), Van Campenhoudt (M.) et Viscogliosi (N.)¹, « Le projet européen DHYDRO : la normalisation à l'épreuve d'un forum terminologique », <http://www.loria.fr/projets/MLIS/DHYDRO/public/publiAETdhydro.pdf> ;
3. GENETRIX, <http://www.genetrix.org/genetrixml/index2-fr.html> ;
4. Gouadec, D., « De la formation à l'assistance. Pour le développement des pratiques terminographiques », IV Journées Internationales de Terminologie - Barcelone, 23 et 24 octobre 1997 ;
5. Institut National de Recherche en Informatique et en Automatique, « Projet LeD Langue et Dialogue », <http://www.inria.fr/rapportsactivite/RA2001/led/led.pdf> ;
6. LeBoeuf, P., « Des modèles et des identifiants : les travaux de normalisation (ISO/TC46, AFNOR/CG46) potentiellement profitables au Web sémantique », http://www.aif.asso.fr/system/imaget/stories/lp2002_selection_am.pdf ;
7. Mangeot-Lerebours, M., « Environnements centralisés et distribués pour lexicographes et lexicologues en contexte multilingue », <http://www.clips.imag.fr/qeta/mathieu.mangeot/MM-These/> ;
8. Otman Gabriel, « Les bases de connaissances terminologiques : les banques de terminologie de seconde génération », in *Meta*, XLII, 2, 1997 ;
9. Popescu-Belis, A., (TIM / ETI, Université de Genève), « Normes basées sur XML pour la gestion de ressources terminologiques et lexicales », http://andreipb.free.fr/cours_xml_web/cours13/cours13.pdf ;
10. Roberts, Roda P. et Langlois, Lucie, « L'apport de l'informatique à la recherche lexicographique », *Meta*, XLVI, 4, 2001 ;
11. *Projet METAFOR* (Rennes), http://www.uhb.fr/urfist/metafor/metafor_historique.htm ;
12. Romary, Laurent, « Un modèle abstrait pour la représentation de terminologies multilingues informatisées TMF Terminological Markup Framework », in *Cahiers GUTenberg* no 3940— Mai 2001 <http://www.gutenberg.eu.org/pub/GUTenberg/publicationsPDF/39-romary.pdf> ;
13. Romary, L., Van Campenhoudt, M., « Normalisation des échanges de données en terminologie : le cas des relations dites *conceptuelles* », Conférence TIA-2001, Nancy, 3 et 4 mai 2001, <http://www.termosciences.fr/IMG/pdf/RomaryCampenhoudt.pdf> ;
14. XXX, *Gestion de données terminologiques : principes, modèles, méthodes* http://www.atilf.fr/perso/salmon-alf/telechargement/Chapitre_Hermes.pdf .