L'informatisation du ThEMA

Indépendance, liberté et multilinguisme

Marjorie Burghart

Édition électronique

URL : <http://journals.openedition.org/ccrh/3025>

DOI : 10.4000/ccrh.3025

ISSN : 1760-7906

Éditeur

Centre de recherches historiques - EHESS

Édition imprimée

Date de publication : 10 janvier 2005

ISSN : 0990-9141

Référence électronique

Marjorie Burghart, « L'informatisation du ThEMA », *Les Cahiers du Centre de Recherches Historiques* [En ligne], 35 | 2005, mis en ligne le 24 mai 2011, consulté le 08 mai 2019. URL : <http://journals.openedition.org/ccrh/3025> ; DOI : 10.4000/ccrh.3025

Ce document a été généré automatiquement le 8 mai 2019.

Article L.111-1 du Code de la propriété intellectuelle.

L'informatisation du ThEMA

Indépendance, liberté et multilinguisme

Marjorie Burghart

- 1 La forme de l'informatisation du ThEMA fut liée aux choix d'un développement académique indépendant. Il s'agissait d'être maître de l'outil, en conservant la possibilité de faire évoluer librement la base, en toute autonomie ; il convenait aussi d'être maître des données, pour envisager diverses possibilités : décliner le ThEMA sous différentes formes (CDRom ; papier) ; favoriser son intégration à d'autres projets (Thesaurus des sermons médiévaux, par exemple).
- 2 Le choix de l'autonomie passait par la question des ressources humaines et matérielles nécessaires. D'où une orientation vers des logiciels libres (PHP, MySQL), selon une philosophie qui était en accord avec les principes premiers de la recherche fondamentale. De plus ces logiciels empruntaient à des technologies ouvertes, bien implantées, pour lesquelles il était facile de mobiliser des compétences (ce qui est un gage de pérennité, de facilité de maintenance).
- 3 La conception de l'interface devait faciliter l'accès rapide aux informations. Des formulaires de recherche donnent accès à un index des mots-clés ou à une saisie libre ; par des menus déroulants le chercheur visualise des listes contrôlées (recueils, répertoires). La liste des réponses se lit en un tableau synthétique (références , résumé ; mots-clés). Il est possible de rebondir sur les mots-clés. L'accès aux fiches individuelles des récits exemplaires se fait en cliquant sur leur référence. Un réel multilinguisme (anglais, italien, allemand, espagnol, français) a été mis en œuvre dans l'interrogation, et pas seulement dans l'interface de recherche. Le dictionnaire des mots-clés a été traduit et uniformisé dans les cinq langues. Ce dictionnaire (environ 2500 termes) a été mis à profit pour permettre la traduction dynamique des requêtes. Lorsqu'une requête est postée depuis l'interface en anglais, italien, espagnol ou allemand, les mots saisis sont traduits à l'aide du dictionnaire interne. Ainsi le résumé (en français, ou en une autre langue) peut être atteint dans chacune des cinq langues.
- 4 Pour faciliter le travail des indexateurs, et pour gérer au mieux le dépouillement des recueils, la solution suivante a été adoptée : les chercheurs qui indexent un nouveau

recueil disposent d'un formulaire simple (sous « Access ») qui leur permet de saisir les mots-clés de façon contrôlée : en ayant accès au dictionnaire (en cinq langues) en cours de saisie, les doublons et les variations dans la traduction sont évités. Les indexateurs conservent néanmoins la possibilité d'ajouter un nouveau terme et ses traductions, s'ils n'en ont trouvé aucun qui leur convienne. L'administrateur du ThEMA dispose d'une base « Access » liée aux tables MySQL en ligne, via ODBC (*Open DataBase Connectivity*) ; il s'agit d'un format défini par Microsoft permettant la communication entre des bases de données fonctionnant sous Windows et d'autres bases. Lors de la réception d'un nouveau recueil, l'administrateur du ThEMA a la possibilité de contrôler les mots-clés créés et d'en valider la création. Il peut ensuite transférer automatiquement les nouvelles fiches d'*exempla* vers la base MySQL.

AUTEUR

MARJORIE BURGHART

Université Lyon II/Maison de l'Orient