

Quel moteur de recherche pour le métaportail ADONIS ?

Un TGE, les SHS et ADONIS

Le TGE (très grand équipement du CNRS) ADONIS a pour mission de réaliser, pour les sciences humaines et sociales (SHS), une infrastructure permettant un accès unifié aux données et documents. Étant donnée la très grande diversité de ces sciences, de leurs données primaires (textes, images, manuscrits, carnets de fouilles, films, statistiques, enquêtes, etc.), de leurs méthodes variées, de leurs documents publiés, édités (près de 2 000 revues françaises !), de leurs communautés spécialisées, c'est une gageure d'une grande complexité que de concevoir et réaliser une telle infrastructure.

Le TGE ADONIS, depuis deux ans, a pris comme initiative de s'appuyer d'une part sur les opérateurs existants des humanités numériques, que ce soit dans le domaine du stockage, de l'archivage, de l'hébergement, du signalement – dont les centres de ressources numériques du CNRS –, de la documentation ou de l'édition. Bien sûr, l'ABES est l'un de ces opérateurs principaux ; de même CLEO, qui avait développé le moteur de recherche *in extenso*.

D'autre part, cette initiative mobilise, autant que faire se peut, les utilisateurs et leurs institutions universitaires ou de recherche, utilisateurs innovants et experts pour piloter et « prototyper » dans un premier temps les meilleures réalisations.

À la convergence de cette double initiative, le TGE ADONIS prend en compte tous les services web possibles comme toutes les données, ressources et publications des SHS mises en ligne ou numérisées. C'est donc tout un ensemble fort riche mais hétérogène de sites de services ou de ressources qu'un métaportail doit intégrer.

Celui-ci devra apporter non seulement des informations pertinentes, mais également interconnecter des données primaires et secondaires, afin que l'utilisateur puisse y trouver le « compagnon » qui contextualise efficacement ses recherches. Côté opérateur, il faudra que ce métaportail apparaisse comme une réelle valeur ajoutée d'ensemble, indépendamment des services particuliers ou des lignes éditoriales spé-

cifiques de chacun des portails disciplinaires ou thématiques.

Depuis 10 mois le TGE ADONIS, aidé de la société ATOS Consulting, a mené à bien la définition technico-fonctionnelle d'un tel métaportail. Au centre et en priorité, devront être réalisés cette année, par un premier marché, un moteur de recherche et un socle de connectivité, qui permettront à cette plateforme d'évoluer au fur et à mesure que des opérateurs de services accepteront de laisser indexer leurs ressources ; au fur et à mesure également que les utilisateurs « producteurs » de données, chercheurs et enseignants-chercheurs des SHS, exposeront leurs données et métadonnées.

Une récente étude américaine sur les cyber-infrastructures de recherche dans le domaine des *Digital humanities** faisait ressortir quatre exigences majeures à surmonter :

- maîtriser l'échelle de croissance des données numériques ;
- mobiliser les outils linguistiques et de communication ;
- caractériser les relations spatio-temporelles ;
- intégrer les réseaux sociaux.

C'est, en substance, un tel programme que s'est donné le TGE ADONIS pour ce moteur de recherche, et pour les différents « lots » de réalisation qui viendront se « connecter » au fur et à mesure de la montée en charge.

Il s'agit bien sûr d'un côté d'indexer un nombre exponentiel de données et métadonnées, mais dans un souci de présentation de différents services pouvant être apportés, par filtrage ou cadrage successif au sein du métaportail. Pour ce faire, le moteur doit pouvoir repérer les données identifiées au sein du maximum de référentiels ou thesaurus déjà élaborés au sein de la recherche, dans les domaines documentaires ou bibliothéconomiques, qu'il s'agisse de données numérisées ou issues du web lui-même.

Il est certain que la valeur ajoutée d'un tel moteur dépendra en second des performances linguistiques et statistiques, élément important du cahier des charges,

dans le plein-texte comme pour les données structurées. À cet égard, la maîtrise des indexations nominales et des croisements bibliographiques sera une priorité. Une grande attente des sciences humaines et sociales réside dans la mise en perspective pertinente en terme de géolocalisation, mais aussi de « chronolocalisation » des données et documents. De nombreuses approches de « visualisation » de ces coordonnées doivent pouvoir enrichir les services proposés.

Enfin, la dimension collaborative des enseignants-chercheurs, étudiants et spécialistes des humanités, doit être utilisée pour administrer et personnaliser les services, afin d'enrichir les communautés en terme de réseaux sociaux.

Si les humanités numériques savent se doter d'instruments technologiques d'une telle ambition, c'est à la fois la richesse sémantique des contenus et la complexité des réseaux collaboratifs qui pourront se développer, au profit des SHS, en France comme à l'échelle internationale.

Y. Maignien

 yannick.maignien@tge-adonis.fr

TGE ADONIS
CNRS - UPS 2916
 <http://www.tge-adonis.fr>
Yannick Maignien, directeur
du TGE ADONIS
 01 55 42 83 10  83 19
ADONIS  contact@tge-adonis.fr
 4 rue Lhomond 75005 PARIS

* *Working Together or Apart : Promoting the Next Generation of Digital Scholarship*
Report of a Workshop Cosponsored by the Council on Library and Information Resources and The National Endowment for the Humanities, March 2009
<http://www.clir.org/pubs/abstract/pub145abst.html>