

TERoogle21? Vers des moteurs thématiques territoriaux

L'ATELIER EST ANIMÉ PAR Jean DELAHOUSSE, de MONDECA

INTERVENANTS : Nathalie PEYPELUT, Fabrice LACROIX, Jean-Paul LEOTARD, Michel GIRAN

Présentation des intervenants



FABRICE LACROIX

PRÉSIDENT D'ANTIDOT



JEAN DELAHOUSSE

PRÉSIDENT D'ANTIDOT



JEAN-PAUL LEOTARD

CHARGÉ DE LA COOPÉRATION
INTERNATIONALE AU MINISTÈRE
DU TOURISME



MICHEL GIRAN

FONDATEUR D'ADOME - PLANÈTE
ÉCOLOGIE

NATHALIE PEYPELUT, Division services de THALES / TerraData: Je suis responsable des études avancées chez THALES et je pilote le projet TerraData sur le thème du patrimoine numérique au sein d'un pôle de compétitivité en Ile-de-France. Pour les PME et les grands groupes adossés à ce pôle, l'objectif est de **construire un projet coopératif sur la valorisation multimédia de la numérisation des territoires et de leurs ressources.**

FABRICE LACROIX, Président d'ANTIDOT: Je suis le fondateur d'ANTIDOT, un éditeur de logiciels spécialisé dans les moteurs de recherche et dans les solutions d'accès à l'information : nous travaillons beaucoup sur les moteurs sectoriels (ou moteurs verticaux).

JEAN DELAHOUSSE, MONDECA: Je dirige MONDECA, une société spécialiste du web sémantique.

J E A N - P A U L L E O T A R D , **Chargé de la coopération internationale au ministère du Tourisme** : Je suis chargé de la coopération internationale au ministère du Tourisme et impliqué dans des projets faisant appel aux nouvelles technologies.

M I C H E L G I R A N , **Fondateur d'Adome – Planète Ecologie** : Je suis co-organisateur de TIC21. L'Internet me fascine dans le sens où il matérialise la « toile de la vie ». En effet, d'après des auteurs comme Fritjof Capra, tous les systèmes vivants, que l'on en soit conscient ou non, sont reliés entre eux dans une sorte de grande toile de vie. Voilà pourquoi il me semblait évident que le monde du vivant se saisisse de ce formidable outil qu'est Internet.

10 ans après sa création, ADOME voit enfin aboutir un projet baptisé « Ecobase 21 », une sorte de « GOOGLE du développement durable » relié à des réseaux de connaissance humaine. L'intervention humaine est indispensable pour ajouter du sens, l'informatique n'est qu'un outil.

Quelques enjeux de l'accès à l'information

J E A N - P A U L L E O T A R D : Le tourisme peut contribuer à la réduction de la pauvreté dans le monde, cette réduction étant l'un des objectifs du millénaire. L'enjeu pour les pays en voie de développement est de proposer **une offre touristique fondée sur de petits opérateurs**, et non sur des tours opérateurs accaparant la grande part de la valeur ajoutée. Comment, pour ces pays, rendre lisible et vendable cette offre sur Internet ?

M I C H E L G I R A N : Sur la thématique du développement durable, comment organiser l'information et trouver un mode de travail commun alors que cette thématique est par définition transdisciplinaire ? L'enjeu est de tisser une toile sémantique entre des acteurs aussi différents que les institutionnels, les ONG, les associations, les professionnels et les citoyens. Si quelqu'un cherche par exemple une information sur les « espèces menacées », le moteur doit lui permettre d'aller plus loin et de préciser sa question : La question posée a-t-elle un lien avec des notions proches telles que : écosystèmes, biodiversité, diversité biologique, WWF, etc. ? Notre moteur permet de guider le questionneur vers des réponses très précises car nos collaborateurs ont identifié un lien de sens entre espèces menacées, écosystèmes, biodiversité, diversité biologique, WWF, etc. et ces liens de sens sont « rentrés » par nos soins dans le moteur.

Les différents modes d'accès à l'information

F A B R I C E L A C R O I X : Les moteurs de recherche donnent accès à l'information selon plusieurs modes.

Les textes et les mots

Google illustre bien l'offre de ces moteurs de recherche basés sur des textes et des mots. Cette offre très basique donne une vision du monde « documentaire ». L'utilisateur se voit proposer l'exploration de millions de pages à chaque recherche : bien qu'émerveillé (pour l'instant) par cette étendue, l'utilisateur ne ressort pas grandi de cette démarche. **Cette abondance de résultats serait plutôt un aveu d'échec pour les nouvelles technologies.**

Les plans de classement (ou taxonomies)

Au lieu de chercher une épingle dans un amoncellement d'informations, une autre stratégie consiste à trier l'information et à l'organiser pour affiner la recherche. C'est une vision héritée du siècle des encyclopédistes, qui structurent et classent rigoureusement l'information (classement du règne animal par exemple). Ces plans de classement sont bien maîtrisés dans les domaines scientifiques. Pour autant, ils restent très délicats à obtenir, car la représentation de la connaissance est elle-même un parti pris. Cette stratégie est donc réservée à des champs de savoirs homogènes et restreints, pour lesquels de petites communautés professionnelles peuvent aboutir à un consensus. S'agissant des connaissances générales, on ne peut mettre toute la planète autour de la table !

“ Cette stratégie est donc réservée à des champs de savoirs homogènes et restreints ”

J E A N D E L A H O U S S E : Sur des thématiques resserrées, des liens entre les sujets sont possibles.

Les bases de connaissances et les liens entre les sujets

Nous travaillons par exemple sur de tels outils avec la Fédération nationale des parcs naturels régionaux (FNPNR). Dans ces parcs, on traite de « chartes », d'« accords avec les agriculteurs », de « biodiversité ». Dans le cas de la FNPNR, il a été possible de définir l'ensemble des sujets d'intérêt pour les parcs et de les relier entre eux. Une fois ce maillage organisé, l'accès au contenu de l'information est possible. Dans ce cas encore, cette démarche n'est possible que si le consensus existe au sein de la communauté concernée (pour un Agenda 21 à l'échelle d'une ville par exemple).

Les modes d'accès à l'information par les textes, les plans de classement et les bases de connaissances font appel tous les trois à la représentation texte : d'autres modes font appel à la 3D.

Les types de représentation 3D : zoom sur TerraData

N A T H A L I E P E Y P E L U T : Avec TerraData, nous souhaitons en effet rendre les informations pertinentes grâce à la 3D. Les enjeux économiques et culturels sont majeurs pour la société de l'information et la création de contenus. En associant des informations pertinentes et en les représentant dans l'espace, les SIG 3D aident à communiquer, permettent d'obtenir des simulations, et aident à la décision des collectivités locales. Ces types de représentation trouvent leur application dans de nombreux domaines :

- les services d'informations urbains et le nomadisme
- les services d'informations culturelles et touristiques
- la production de contenus pour les applications éducatives et ludiques
- la muséographie multimédia et les installations artistiques interactives
- l'entraînement à la conduite et la simulation des risques
- les systèmes de sécurité civile et militaire.

Le projet TerraData associe sur le sol français des PME, des grands groupes et 9 laboratoires qui ont ensemble une véritable carte à jouer sur ces domaines pour contrecarrer les offensives américaines.

J E A N D E L A H O U S S E : En résumé, l'approche aujourd'hui popularisée par Google va bel et bien évoluer vers une vision plus organisée de l'information, et des solutions variées existent déjà aujourd'hui. On peut aussi penser que ces évolutions seront liées à celles du téléphone mobile, une interface nettement plus démocratique que le PC.

Un effort d'organisation des contenus

Thesaurus

J E A N - P A U L L E O T A R D : La brique indispensable à tout mode d'organisation de l'information est le « thesaurus » : il s'agit d'un arbre de termes définissant les « tiroirs » dans lesquels viendra se ranger l'information. Ce thesaurus permet aux différents acteurs d'un secteur d'activité de parler le même langage. A condition que le thesaurus soit traduit (et c'est le cas dans le secteur du tourisme¹), ce thesaurus permet même de monter des systèmes inter-opérables à l'échelle internationale.

“ Ce thesaurus permet aux différents acteurs d'un secteur d'activité de parler le même langage ”

J E A N D E L A H O U S S E : D'abord construit pour les besoins de la documentation, le thesaurus du tourisme a été traduit à l'international. Les investissements sur cette infrastructure de connaissances (relativement lourds) ne peuvent être financés que par les pouvoirs publics. Aux acteurs locaux est ensuite laissé le soin d'offrir les services en utilisant l'infrastructure existante.

J E A N - P A U L L E O T A R D : A l'origine, le thesaurus est en effet né du besoin des collectivités françaises d'échanger leurs informations touristiques sur un format commun : derrière le terme « hébergement », chacun sait ce qui se cache grâce au thesaurus. Certains territoires sont dorénavant passés au format commun sous XML. Et désormais, un Finnois désirant se rendre dans le Sud-Ouest peut interroger en finnois une association locale sur les mots « foie gras, musique baroque, ski de fond » et recevoir les réponses en finnois. Le tourisme local s'en porte bien et reçoit étrangement beaucoup de Finnois...

D E L A S A L L E : Comment appréhender l'actualisation de ces bases de connaissances ? Les utilisateurs peuvent-ils apporter leur contribution ?

J E A N D E L A H O U S S E : De nouvelles activités touristiques naissent et les événementiels culturels sont différents chaque semaine de la saison : l'association locale prend en charge l'actualisation de ces données. Mais ce qui reste possible dans le secteur du tourisme sera moins évident dans un secteur vivant de fortes évolutions technologiques comme les télécommunications.

Le projet GEMET

M I C H E L G I R A N : J'aimerais détailler l'exemple du thesaurus utilisé par l'agence européenne de l'environnement. L'une des missions de cette agence consiste à faire remonter et à consolider les informations environnementales en provenance des Etats membres jusqu'à l'échelon européen en vue d'orienter les politiques européennes.

Depuis la signature par des pays européens des conventions de Rio et d'Aarhus, qui font obligation aux pouvoirs publics de donner libre accès à toute information environnementale,

¹ Le thesaurus pour le secteur du tourisme a été récemment traduit en 5 langues et adopté à l'échelle internationale.



l'agence européenne cherche à enclencher un processus de restitution de l'information aux citoyens. Elle a donc travaillé sur un thesaurus fondé nommé GEMET (environ 10 000 termes traduits dans toutes les langues). Grâce à ce thesaurus, il est facile de faire par exemple correspondre la sous rubrique « déchets hospitaliers » en « France » avec la sous rubrique « déchets hospitaliers » en Italie, en Allemagne, en Angleterre... ce qui permet de comparer les pratiques en matière de déchets hospitaliers dans ces pays Des organisations internationales ont à leur tour décidé de structurer leur information selon cette architecture et l'objectif de l'agence européenne est maintenant d'en faire une norme. En tout cas, aujourd'hui, ce thesaurus est de plus en plus partagé et ouvert. On peut penser qu'il couvrira un jour la thématique du développement durable dans son ensemble et pas seulement celle de l'environnement.

Tout n'est pas simple pour autant dès lors qu'on a besoin de traductions : dans les langues asiatiques par exemple, il n'existe pas de mots ayant le même sens que notre mot nature, car cette dernière n'est tout simplement pas considérée comme un sujet d'étude extérieur à nous-même : nous n'avons pas la même vision du monde !..

FABRICE LACROIX : Précisons la différence entre un thesaurus et une encyclopédie comme Wikipédia². Le thesaurus est simplement la colonne vertébrale, le squelette auquel sera rattachée ensuite l'information, tandis que Wikipédia présente déjà un contenu.

JEAN DELAHOUSSE : Dans le secteur du tourisme par exemple, chaque territoire pourra parler le même vocabulaire en utilisant le thesaurus, mais les contenus sur l'information touristique seront évidemment différents pour chaque territoire.

Thesaurus du tourisme

JEAN-PAUL LEOTARD : Pour éviter la distribution de masse des produits touristiques, nous structurons avec les acteurs locaux une offre diffusable par Internet. Cette démarche requiert évidemment une bonne organisation entre les acteurs du tourisme, le respect de chartes, la mise en place d'agréments... Ce combat mené pour leur indépendance est exactement le même que celui de Max Havelaar dans le commerce éthique.

Vous avez probablement entendu parler du phénomène El Nino, à qui l'on doit, entre autres, l'augmentation de la température de l'eau aux Philippines. L'ampleur est telle que beaucoup de pêcheurs doivent cesser leur activité suite à la raréfaction des ressources. Avec d'autres partenaires à l'échelle européenne, nous avons aidé ces pêcheurs à se convertir sur une activité de centre de plongée. Nous avons simplement récupéré et envoyé les listing des centres de plongée de toute l'Union européenne (cette activité est très organisée car très réglementée) : aujourd'hui, les Philippines ne savent plus comment répondre à la demande. Voilà ce que permettent aujourd'hui les bases de connaissances et l'Internet.

Le projet Ecobase 21

MICHEL GIRAN : Google ne permettant pas d'apporter du sens à la recherche, nous avons voulu, en menant le projet Ecobase 21 (avec ANTIDOT et MONDECA), travailler sur le sens et les contenus.

Le moteur sectoriel « Ecobase 21 » se présente comme une base de données puisant dans 1 500 fiches organisées suivant une toile de sens (grâce à des mots-clés secondaires).

[Démonstration en temps réel au cours de l'atelier sur le site www.ecobase21.antidot.net : une recherche sur le mot « recyclage » fait apparaître des sous-menus pollution /

² Wikipédia est un projet d'encyclopédie libre, gratuite et multilingue que chacun peut enrichir : www.wikipedia.fr

propreté urbaine / parties prenantes. La recherche peut s'affiner jusqu'aux fiches qui comportent elles-mêmes plus de 35 000 liens vers des sites web sélectionnés]

Une telle toile de sens, mettant au même niveau par exemple « développement durable » et « décroissance », serait éminemment politique. Voilà qui fait naître immédiatement la question : quels seront les spécialistes qui tisseront cette toile pour les moteurs thématiques, des institutionnels, des ONG, des universitaires ? Quelle vision du monde ont-ils et quelle légitimité pour la proposer ?



Echanges avec la salle

DE LA SALLE : Que penser de l'hégémonie de Google et de la volonté de faire émerger un challenger européen ?

FABRICE LACROIX : Nul ne pourra aller contre cette réalité : Google a recueilli une adhésion large, les utilisateurs ne vont pas soudainement modifier leurs usages simplement parce qu'est annoncée la naissance d'un moteur européen. Ne croyons pas que ce projet sauvera l'Europe.

Ce qui est sûr, c'est qu'aujourd'hui Google a cessé d'investir dans la technologie : il a trouvé son modèle économique, d'ailleurs très robuste. Mais la place est libre pour des moteurs couvrant un périmètre et une cible définis, et appliquant des techniques similaires à Ecobase 21.

DE LA SALLE : Mais votre société s'est-elle récemment lancée ? De quoi vivez-vous ?

FABRICE LACROIX : Nous existons depuis plusieurs années et vivons de clients, pas de subventions !..

JEAN DELAHOUSSE : Actuellement, le marché est porteur et les sociétés dans ce domaine sont en croissance.

DE LA SALLE : Comment combiner la cartographie aux données « textes » ?

NATHALIE PEYPELUT : L'idée de TerraData est justement de travailler sur cet assemblage. Prenons l'exemple de la ville de Cannes, dont les 22 000 bâtiments répartis sur 21 Km² ont été modélisés et restitués en 3D. Cette réalisation – coûteuse – est certes magni-

fique mais totalement décorrélée de toute l'information disponible sur les activités culturelles de la ville de Cannes, sur la circulation... Or, il est aujourd'hui possible de combiner ces informations, de les « mapper » sur l'information géographique, et de faire de ce bel objet un véritable outil de communication. L'objectif du projet TerraData est de faire diminuer les coûts de modélisation et d'appréhender la recomposition des informations.

DE LA SALLE : Comment le responsable d'un Agenda 21 peut-il se saisir de l'outil Ecobase ?

FABRICE LACROIX : Ecobase 21 illustre notre propos aujourd'hui, mais est encore un démonstrateur expérimental. Il faut désormais animer l'information, faire en sorte qu'une communauté prenne en charge les outils d'actualisation.

DE LA SALLE : Comment rendre cet outil « vendable » ?

MICHEL GIRAN : Le modèle économique à trouver ne sera sans doute pas le « tout gratuit » ni le « tout payant ». Ecobase 21 se déclinera probablement en plusieurs entrées selon l'utilisateur (entreprise, collectivité, association...). On peut imaginer un accès gratuit pour l'utilisateur jusqu'à ce qu'il entre dans l'information très spécifique.

DE LA SALLE (CATHERINE TRAUTMANN) : Envisagez-vous une offre de services venant se greffer aux moteurs de recherche ? Quels seraient les offreurs : les acteurs ayant conçu le moteur ou certains utilisateurs connectés au moteur ?

FABRICE LACROIX : Nous ne sommes que fournisseurs de technologies, et n'avons pas vocation à devenir opérateurs de solutions. Nous sommes des facilitateurs.

DE LA SALLE : Faut-il rétribuer les contributeurs lorsqu'ils enrichissent le contenu des moteurs de recherche ?

FABRICE LACROIX : Les contributeurs à Wikipédia sont avant tout des passionnés : ils n'enrichissent pas l'encyclopédie pour gagner des crédits de contribution ou des points. Ils y voient la possibilité d'avoir une tribune, le caractère participatif de Wikipédia les mobilise.

JEAN DELAHOUSSE : Tout dépend de la communauté en question : Wikipédia fonctionne grâce à la passion ; le projet avec la FNNP a dû être animé par des professionnels, une contrepartie financière doit être envisagée, etc...

MICHEL GIRAN : Techniquement, les outils sont conçus pour être enrichis par des contributions. Mais qui contribue ? Pour quel public ? C'est plutôt sur ces questions qu'il nous faut réfléchir et choisir.

DE LA SALLE : A des outils comme Ecobase 21, les communautés professionnelles aimeraient voir s'accoler des bases de connaissances qui leur sont propres, des outils de mise en œuvre, de suivi opérationnel, de contrôle... qui leur donneraient une véritable capacité de diagnostic et de décision.

DE LA SALLE : Les outils de recherche comme Kartoo font appel à la cartographie, allez-vous y passer ?

FABRICE LACROIX : Kartoo a en effet adopté ce mode de représentation que l'on peut très bien décliner pour les moteurs que nous développons : nous disposons déjà des interfaces.

DE LA SALLE : Comment intégrer les connaissances des utilisateurs dans TerraData ?

NATHALIE PEYPELUT : A ce stade du projet TerraData, nous en sommes à la recherche sur la technologie. Les questions sur les perspectives collaboratives viendront dans un second temps. De toute façon, l'intégration de la visualisation sur divers terminaux ouvrira aux utilisateurs des perspectives encore inconnues aujourd'hui : bien souvent, en réduisant le coût

d'une technologie, les utilisateurs se l'approprient d'une manière bien différente de celle prévue à l'origine.

DE LA SALLE (CATHERINE TRAUTMANN) : Je remarque que tout projet R&D dans ces domaines doit être conduit avec une population jeune, c'est elle qui de toute façon va utiliser les résultats, et qui est plus habituée à se mouvoir dans l'espace et la 3D (grâce aux jeux vidéo).

Par ailleurs, la dimension démocratique est essentielle si l'on souhaite que ces projets recueillent une adhésion la plus large possible. Bien souvent, en tant qu'élue, j'observe que l'on se précipite dans le projet et que l'on se demande ensuite comment le partager... De manière générale, faire partager les travaux de recherche devient de plus en plus crucial.

NATHALIE PEYPELUT : C'est ce qui me conduit à penser que l'e-administration met aujourd'hui de côté l'e-citoyen. De nombreux projets dans le domaine de la e-administration voient le jour sans que les citoyens aient été interrogés au préalable.

DE LA SALLE : JJe mentionne l'existence d'un projet européen qui consiste à représenter en 3D les informations portant sur une commune rurale. Grâce à cet outil, un agriculteur, par exemple, peut visualiser et mieux comprendre son impact sur l'environnement, il peut simuler des scénarios, visualiser les effets, comparer plusieurs scénarios...

FABRICE LACROIX : Les nouvelles générations ont l'habitude des blogs, participent à des forums, s'envoient des SMS. Elles sont nées en même temps que ces outils d'expression et vont se les approprier plus rapidement que nous.



Sigles

F P N R F

Fédération des parcs naturels régionaux de France

O N G

Organisation non gouvernementale

S I G

Système d'information géographique