

Interopérabilité et Services Web standardisés du "Human to Human" au "Machine to Machine"

Henri Pornon, animateur du groupe OGC AFIGEO

SIGLL 2011 - Session Interopérabilité : 18 mai 2011

Mise en contexte

- **De nombreux dispositifs de mutualisation de l'information géographique**
 - Infrastructures de données géographiques
 - Observatoires territoriaux (économie, biodiversité, énergie, etc)
- **Ils s'appuient sur les standards d'interopérabilité de l'OGC**
 - Interopérabilité des plates-formes avec les adhérents / partenaires
 - Interopérabilité des plates-formes entre elles
- **Ces standards permettent notamment à des humains et des machines :**
 - De moissonner des métadonnées
 - De consommer des services Web donnant accès à des données et parfois des traitements associés

Contexte	Interaction via les métadonnées
Human to Human (H2H)	Une personne après avoir identifié des métadonnées, téléphone à leur gestionnaire pour obtenir plus d'informations sur ces données et pouvoir y accéder.
Human Computer Interaction (HCI)	<p>Une personne exécute une requête de recherche de métadonnées dans un logiciel</p> <p>Le logiciel présente les métadonnées, triées à l'aide de certains algorithmes, à l'utilisateur</p> <p>La personne active les liens permettant d'ouvrir les fichiers de données</p> <p>Ces interactions sont de plus en plus gérées par des applications Web (technologies du WWW), et sont souvent présentées comme des « Services » (sur le Web). Il va sans dire que ce terme apporte une confusion dans la bonne compréhension de ce que sont et font par ailleurs les « Web Services »</p>
Machine to Machine (M2M)	Un logiciel envoie des requêtes de recherche ou de moissonnage de métadonnées à d'autres logiciels et indexe les résultats

Contexte	Interaction via les « Web Services »
Human to Human (H2H)	Les « Web Services » sont par définition des interfaces programmatiques
Human Computer Interaction (HCI)	<p>Une personne configure les paramètres d'accès à un « Web Service » WMS pour l'utiliser dans son SIG bureautique habituel.</p> <p>Le logiciel présente à un utilisateur une carte affichant le résultat de l'appel du « Web Service »</p>
Machine 2 Machine (M2M)	Un logiciel orchestre un chaînage de « Web Service » pour réaliser une fonction à valeur ajoutée

- Ne jamais oublier que les Services Web ont comme premier objectif de faire communiquer les machines entre elles !

Des problèmes ?

- Les standards OGC
 - ne résolvent pas tous les problèmes : ceux qui restent non résolus
 - posent de nouveaux problèmes : ceux qui émergent de l'utilisation des standards
- Leur mise en œuvre permet d'identifier des problèmes
 - Identification des données et services
 - Qualification et sélection
 - utilisation
- Qui se posent de façon différente, suivant l'utilisateur :
 - Humain ou Machine

Identification

- Humain
 - Comment savoir où aller chercher des métadonnées (méta-annuaires, centralisation des métadonnées ?)
 - On ne cherche que dans des catalogues que l'on connaît
- Machine
 - Obligation de fournir à l'ordinateur une liste de catalogues, donc de prédéfinir les recherches
 - Pas de services de méta-annuaires incluant des critères thématiques ou géographiques

Sélection (1)

- Comment faire pour qu'une requête dans le catalogue renvoie le plus de données pertinentes possible (et si possible, ne renvoie que des données pertinentes)
- Humain
 - Capable d'exploiter une certaine ambiguïté des concepts (PLU, plan d'urbanisme, plan de zonage, etc)
 - Plus les mots clés sont spécifiques et diversifiés, plus les résultats de ses requêtes seront pertinents
 - Plus les mots clés sont génériques, plus il aura de résultats non pertinents
- Machine
 - Ne peut exploiter que des concepts et des mots clés précis
 - Plus les mots clés sont spécifiques et diversifiés, moins il obtiendra de résultats
 - Plus les mots clés sont génériques, plus les métadonnées seront accessibles
- Suffit-il d'enrichir les thésaurus et ontologies et d'interdire les mots clés hors thésaurus ?



Sélection (2)

- Problème des emprises (bounding boxes)
 - Humain
 - Sait s'accommoder de l'imprécision de la localisation
 - Machine
 - Ne peut éliminer tout seul des emprises inadaptées qui intersectent son territoire de recherche
- Problème des recherches multilingues
 - Encodage des thésaurus ou tables de correspondances entre concepts (ce qui nécessite qu'il y ait un bon niveau de correspondance sémantique)
 - Problèmes de terminologie et de sémantique



Qualification de la sélection

- Humain
 - Il utilisera les zones de texte libre "Résumé" et "Généalogie)
 - L'examen des diverses métadonnées collectées peut-être long et fastidieux mais il dispose de moyens cognitifs pour distinguer les données et choisir celles qui lui conviennent
- Machine
 - Rarement en capacité de comparer les services et de choisir automatiquement le plus pertinent
 - Il ne peut choisir que sur des critères précis pré-codés : emprise ou nom de territoire, mots clés issus de thésaurus
 - Il pourrait éventuellement se baser sur des critères qualités renseignés de façon homogène (ce n'est jamais le cas)
- Et comment savoir si divers services ou lots de données sont homogènes et peuvent être assemblés ?



Utilisation

- Comment assembler divers services et lots de données ?
 - Humain peut définir une configuration à la carte, service par service et homogénéiser la présentation des données
 - Machine ne peut résoudre ce problème sans intervention humaine que si les données sont homogènes et si des critères de présentation (SLD) ont été associés aux services



Des problèmes d'interopérabilité à plusieurs niveaux

- Problèmes techniques
- Problèmes d'ordre sémantique
- Problèmes d'ordre linguistique
- Problèmes organisationnels ?
- Les métadonnées sont le plus souvent saisies pour des acteurs humains et sont rarement utilisables en l'état par des systèmes informatiques
- Un niveau supplémentaire de standardisation est nécessaire pour atteindre l'objectif d'interopérabilité entre machines



Conclusion : pistes de réflexion

- Pistes concernant les moyens actuels
 - Imposer des ontologies ?
 - Les modèles standardisés INSPIRE apportent-ils la solution ?
 - Standardiser le nommage des métadonnées et les mots clés ?
 - Mettre au point de nouveaux standards ?
- Pistes plus prospectives
 - Mettre au point des services de recherche plus élaborés
 - Etudier d'autres moyens techniques utilisés dans des contextes computer to computer : outils de tri de l'intelligence économique par exemple



Merci de votre attention

