

Mardi 16 octobre 2012

Datalift Camp : résumé d'une journée de formation-action au « Web Sémantique » chez Etalab

[Etalab](#), chargée de coordonner l'ouverture des données publiques, a accueilli le 10 octobre dernier « le DataLift Camp » en présence des membres du consortium Datalift, Charles Nepote ([FING](#)) et François Scharffe ([LIRMM](#), [Datalift](#)), et de plus d'une quarantaine de participants producteurs de données, fournisseurs de plateformes, acteurs publics et experts de l'Open Data en France (de l'EURECOM, l'INSEE, l'INRA, Logica, l'IGN, le Ministère de l'Intérieur...).



Après le mot d'accueil de Charles Ruelle pour Etalab, Charles Nepote engage la conversation

La matinée a commencé par une introduction au web sémantique, ses enjeux et ses usages. Charles Ruelle, directeur technique de la mission Etalab, a d'ailleurs rappelé à cette occasion « Le « *web des données* » (ou *web sémantique*), est la promesse d'un web centré sur les données, un web où les données de toutes formes et de tous acteurs se trouvent normalisées, identifiées sans ambiguïté et reliées les unes aux autres. Ces technologies appliquées à l'Open Data faciliteraient la création de nouveaux services, de nouveaux usages, et de nouveaux croisements de données. ».

Les technologies du web sémantique permettent notamment aux données ouvertes « Open Data », d'être normalisées, identifiées sans ambiguïté et d'être reliées les unes aux autres.

Vidéo pédagogique de présentation du web des données : [« Le web des données ouvertes et liées, qu'est ce que c'est ? »](#)



Ont ensuite suivies les présentations de Gabriel Kepeklian, responsable Recherche et Développement à [Atos](#) et de François Scharffe, coordinateur du projet Datalift et chercheur en Intelligence Artificielle et en Web Sémantique.



Présentation de DataLift par François Scharffe, Datalift Camp 9 et 10 octobre 2012, p.9

« Les mains dans le cambouis »

Après cette session plénière, les participants ont été invités à soumettre leurs idées de jeu de données et à former des groupes.

Fanny Maestracci de [Suez Environnement](#) a démarré par l'idée de travailler sur « l'eau », « l'impact sur les collectivités », la « qualité de l'eau dans les communes ». C'est ensuite Eric Cossa de [Picardie Tourisme](#) qui a pris la parole, intéressé par les données aussi bien touristiques, que par des croisements de données géographiques, météorologique, et/ou environnementales.



« Brainstorming » en fin de matinée

Après consultation de chacune des personnes présentes lors de cette journée de formation, les groupes « Prénoms », « INSEE », « Evénements », « Environnement », « Géolocalisation » et « DCAT » ont été définis pour le travail de l'après-midi.

La veille, au cours du Datalift Camp du 9 octobre, les groupes avaient abordé les sujets de croisement de données INSEE avec celles du Code Officiel Géographique ; les données de transport ; le budget des collectivités ; les prénoms ; et le croisement de données Open Food Facts.



Le groupe Environnement (Alimentation/Agronomie) travaille à l'interconnexion des données.

Un bilan positif

Les équipes formées ont été enthousiastes, inventives, ce malgré les quelques problèmes techniques rencontrés (de syntaxe, de conversion des données...). Pour certains d'entre eux, l'utilisation de la plateforme Datalift était une première qui s'est avérée prometteuse. Tandis que pour d'autres, participer au Datalift Camp correspondait à une continuité dans leurs recherches :

Par exemple, Christian Quest d'[Open Street Map](#) affirmait l'atout d'une formation au web sémantique : « le web sémantique est avant tout un apport de qualité à un travail de recherche déjà en cours ».

Stéphan Bernard de l'[IRSTEA](#) a parlé de web sémantique comme « d'un moyen de trouver des clefs pour identifier les données utiles dans le cadre d'une recherche appliquée » et « de distinguer des données similaires avec des intitulés différents ».

Enfin, Nicolas Chauvat de [Logilab](#) évoquait l'intérêt du Datalift Camp, sa société étant spécialisée dans l'informatique scientifique et le web sémantique.



Datalift Camp : et maintenant ?

Datalift est un logiciel libre, ce qui permettra aux participants de poursuivre leurs travaux avec le concours des référents du consortium pour arriver à un résultat abouti.

Datalift poursuit ses travaux de développement pour faciliter l'élévation sémantique de données.

Enfin, dans le cadre des [Dataconnexions](#), la FING prévoit un hackathon « données 5 étoiles » courant novembre 2012 pour « valoriser les données sémantiques produites ».

Toute l'équipe d'Etalab remercie vivement les participants et les techniciens experts pour leur contribution enrichissante à cette deuxième session du Datalift Camp.