Desgodets, interopérabilité

# XML et XML-TEI garants de l’interopérabilité

Expliquer

# La production de métadonnées

## L’en-tête TEI

Voir le travail mené sur l’en-tête TEI pour ce qui concerne l’exposition de métadonnées.

Décrire le mapping effectué avec Dublin Core et l’initiative Cahier.

Décrire d’autres possibilités de mapping.

## La création d’un entrepôt OAI-PMH

Justifier du fait qu’on y ait renoncé

Expliquer conformité au moissonneur en cours de développement pour Cahier.

# Identifiants et URI

http://www.w3.org/TR/webarch/

Architecture of the World Wide Web, Volume One, W3C Recommendation 15 December 2004

http://w3ctag.github.io/webarch/

D’après la définition qu’en donne le W3C, « le World Wide Web (www, ou simplement Web) est un espace informationnel dans lequel les objets d’intérêt, référencés comme des ressources, sont identifiés par des identifiants globaux appelés Uniform Resource Identifiers (URI). »[[1]](#footnote-1)

L’architecture du web repose donc en grande partie sur la production d’identifiants qui désignent des ressources sur le web et permettent d’accéder à leurs représentations en utilisant le protocole de transfert hypertexte HTTP (Hypertext Transfer Protocol). Repose donc sur les trois éléments de base suivants :

* Identification : Les URIs sont utilisées pour identifier des ressources, ces ressources pouvant faire l’objet d’actualisation périodique.
* Interaction : Les agents web communiquent en utilisant des protocoles standardisés qui permettent l’interaction à travers l’échange de messages conforment à une syntaxe et une sémantique bien définie. En entrant une URI dans une fenêtre de saisie ou en sélectionnant un lien hypertexte, on indique à un navigateur d’effectuer une action de récupération (retrieval) de la ressource identifiée par l’URI. Avec l’exemple d’une page web, le navigateur envoie une requête HTTP GET au serveur via TCP/IP sur le port 80, et le serveur retourne un message contenant ce qu’il a déterminé comme étant la représentation de la ressource au moment où cette représentation a été générée. Il faut noter que cet exemple est spécifique à la navigation d’information hypertexte, d’autres types d’interaction sont possibles, soit au sein de navigateurs soit à l’aide d’autres agents web.
* Formats : La plupart des protocoles utilisés pour la récupération et/ou la soumission de représentation utilisent une séquence d’un ou plusieurs messages qui pris ensemble contiennent un jeu de données représentationnelles et de métadonnées, pour transférer la représentation entre les agents.

Bonne pratique 1 : Identifier avec des URI

Une ressource doit avoir une URI associée si une autre partie souhaite créer un lien hypertexte vers elle ou réfuter des assertions à son sujet, récupérer ou mettre en cache une représentation de celle-ci, incluant toute ou partie de celle-ci par référence à une autre représentation, l’annoter, ou réaliser d’autres opérations sur elle.

Cf. http://www.w3.org/2001/tag/doc/whenToUseGet.html URIs, Addressability, and the use of HTTP GET and POST, TAG FInding, 21 March 2004. Distingue les cas d’usage de GET et POST.

Dans certains cas, il peut être désirable d’interagir de différentes manières avec la ressource identifiée par une URI. Séparation des URIs, des méthodes.

L’architecture du web permet ainsi une séparation des URIs et des méthodes. Un scheme d’URI’s détermine un ensemble de méthodes d’interactions. Par exemple, le scheme d’URI http est défini dans HTTP/1.1 [RFC2616], qui autorise « un ensemble ouvert fini (open-ended) de méthodes et de headers indiquant le but d’une requête. Il est construit sur la discipline de référence fournie par le Uniform Ressouce Identifier (URI)… pour indiquer la ressource à laquelle une méthode doit être appliquée. »

Cool URIs don’t change

http://www.w3.org/Provider/Style/URI

Plusieurs cas de figures envisagés :

* réorganisation du site web, si doit renommer les URI, c’est que celles-ci ont été mal choisies
* Traiter correctement les versions dans les métadonnées.
* Ne pas dépendre de l’organisation des fichiers dans le système de fichiers. Considérer l’espace des URI comme un espace abstrait, parfaitement organisé, puis faire un mapping sur n’importe quelle réalité que vous utilisez pour l’implémenter. Puis l’indiquer au serveur.
* Ne pas dépendre des technologies de rendu (scripts, etc.). Il n’est pas nécessaire d’exposer le mécanisme avec lequel on fait fonctionner le serveur. Sans quoi change dès que change de système.

Comment enlever les extensions

Négociation de contenu. Choisira le fichier le plus adéquat.

* Configurer le serveur pour faire de la négociation de contenu
* Faire référence à l’URI sans l’extension

Tim Berners Lee, Web Architecture from 50,000 feet http://www.w3.org/DesignIssues/Architecture.html

"Cool URIs for the Semantic Web", *Cool URIs for the Semantic Web.*

http://www.w3.org/TR/cooluris/

1. Ian Jacobs, Norman Walsh, *Architecture of the World Wide Web, Volume One*, W3C Recommandation 15 décembre 2004, http://www.w3.org/TR/webarch/, accédé le 8 juin 2013. Ce document est actuellement en cours de remaniement par le W3C : Ian Jacobs, Henry S. Thompson, *Architecture of the World Wide Web*, W3C Editors’ Draft, 7 juin 2013, http://w3ctag.github.io/webarch/, accédé le 8 juin 2013. [↑](#footnote-ref-1)