

Le e-lien, une solution pour l'extraction et le partage de connaissances structurées dans les documents hypertextuels

Gilles Verley* , J.P. Asselin de Beauville**

Laboratoire d'Informatique, 64, avenue Jean Portalis, 37000 - TOURS

*gilles.verley@univ-tours.fr

**jean-pierre.asselin@auf.org

On propose ici une solution pour extraire et partager des connaissances structurées dans des documents hypertextuels . Pour cela, on a examiné quelles peuvent être les différentes motivations qui amènent un auteur à placer un lien hypertexte dans la page qu'il crée. On a montré alors que certaines de ces motivations représentent une connaissance structurée potentielle dont l'exploitation pourrait être utile à une communauté d'auteurs et de lecteurs.

Ces considérations nous ont amenés à élaborer un langage formel permettant de les expliciter. Un lien explicité dans ce langage devient un lien "étendu" (e-lien). Pour mettre en œuvre ce nouveau lien, on a développé le concept de "serveur de e-liens" d'un point de vue théorique et technologique. L'objectif de ce serveur particulier est de récupérer la connaissance formelle placée par les auteurs dans les e-liens afin de la partager de différentes manières. Il est alors possible pour un auteur de poser un e-lien vers une page « cible » dont l'adresse lui est encore inconnue en indiquant simplement la motivation qui a justifié la pose de cet e-lien.

Illustrons sommairement quelques éléments du langage sur un exemple très simple. Le site de Météo-France dispose d'une page dédiée aux prévisions météorologiques et qui est constituée d'une liste de rubriques telles que « les prévisions marines », « les prévisions terrestres », « les prévisions aériennes ». Ces rubriques deviendront naturellement des e-liens de type hiérarchique permettant de les détailler. Voici comment on explicite simplement le e-lien concernant les prévisions marines:

*Les **prévisions marines** sont une spécialisation des **prévisions météorologiques** dont l'url est <http://www.meteo.fr/prevmeteo.xml>*

Cet e-lien fournit des informations du point de vue de la page « appelante » sans préciser l'adresse de la page « appelée » comme dans le cas d'un lien ordinaire. Par contre, il précise la manière dont cette adresse peut être trouvée grâce à des informations sémantiques de haut niveau. L'activation de cet e-lien permet, au travers d'un serveur de e-liens, de fournir à une base de connaissances des informations actualisées de haute valeur sémantique reçues spontanément et qui vont servir à la recherche de la page répondant le mieux au besoin exprimé.

Complémentairement, le serveur d'e-liens pourra fournir aux utilisateurs des informations contextuelles pertinentes car proches sémantiquement du besoin exprimé dans l'e-lien activé. Pour les auteurs, l'utilisation des e-liens à la place des liens traditionnels facilite, grâce aux fonctionnalités du serveur d'e-liens, le développement et la maintenance de sites individuels ou collectifs.

L'ensemble des concepts présentés très sommairement dans cette page ont été mis en œuvre expérimentalement dans le strict cadre des normes http, html et xml. Les éléments technologiques sont disponibles sur simple demande aux auteurs.