

# Présentation de PubliTool

Michel Cubero-Castan  
INSA - Département Génie Electrique et Informatique  
135, Avenue de Rangueil  
31077 Toulouse cedex 4 (France)  
Michel.Cubero-Castan@insa-tlse.fr

## Abstract

Nous présentons un outil pour la publication multi-format de documents XML. Entièrement rédigé en Java, il permet une expression simple d'enchaînements de feuilles de style XSL. Il intègre un accès protégé au niveau de l'élément interne du document, particularité qui lui permet de générer, à partir d'un document unique, plusieurs vues en fonction du profil de l'utilisateur. Il s'utilise soit en tant qu'application Java autonome, soit en tant que servlet pour une intégration dans un serveur http.

Mots-clés: Serveur dynamique, XML, XSL, XSLT, FO, SVG, Java, Servlet

We present a tool for the multi-format publication of XML documents. Written in Java, it allows a simple expression of XSL style sheets chaining. It integrates a protected access at the level of the internal element of the document, characteristic which enables to generate, from a single document, several views according to the profile of the user. It can be used either as a Java application, or as servlet for an integration in a http-server.

Keywords: Dynamic server, XML, XSL, XSLT, FO, SVG, Java, Servlet

Avec le développement des TICs, les enseignants sont amenés à réaliser des documents pédagogiques sous de multiples formes. Cette multiplicité se décline sur trois dimensions. La première dimension *source* prend en compte le type du document: une image, un document réalisé avec tel ou tel outil, ... La deuxième dimension *format* prend en compte le type d'outil utilisé pour visualiser le document: format html, format pdf, ... La troisième dimension *filtre* prend en compte certaines caractéristiques liées à l'utilisateur du document pour en restreindre l'accès: enseignant, étudiant, domaine réseau, ...

Prenons l'exemple d'un support de cours d'initiation au langage de programmation Java contenant des exercices pour les séances de travaux dirigés. En regardant la première dimension, nous constatons que ce document contient des images au format gif, des dessins réalisés avec le logiciel xfig, des animations sous forme d'applets, des programmes sources java, des exemples d'exécution sur unix ou sur windows. Au niveau de la deuxième dimension, nous pourrions avoir les formats html et pdf au choix. Au niveau de la troisième dimension, le document pourra être consulté par le rédacteur du document (il pourra ainsi voir le contenu complet du document, avec des notations sur les différentes mises à

jour), par un enseignant de TD (il pourra ainsi avoir accès aux solutions), par un étudiant de l'école (il ne pourra avoir accès aux solutions qu'après une certaine date), par un anonyme (il n'aura pas accès aux solutions).

En général, cette multiplicité se retrouve au niveau des documents que le réacteur doit gérer. Il n'y a pas un document unique, mais de multiples documents souvent créés en utilisant le sacro-saint *copie-coller*, source d'erreurs lors des mises à jour ultérieures.

PubliTool est un outil de publication de documents xml, adapté à la production de documents pédagogiques. Il permet, à partir d'un document unique XML, de produire différentes versions destinés à des publics différents. Il s'utilise soit en tant qu'application Java autonome, soit en tant que servlet pour une intégration dans un serveur http. Il utilise l'API 1.4 de Java 2 SDK.

## Les trois dimensions d'un document

A travers la présentation de quelques balises, examinons chaque dimension d'un document.

### Dimension source

Cette dimension offre la possibilité d'intégrer, dans le document XML, des documents non XML. Ce n'est pas en soi un aspect très original de PubliTool, puisque cette possibilité se retrouve dans la plupart des documents XML. PubliTool offre toutefois la possibilité de traiter ces éléments non XML par une application extérieure.

La balise `<latex>` permet d'inclure directement du code latex dans le document xml, code qui sera transformé automatiquement en image grâce à un script composé par une feuille de style XSL. Des balises analogues permettent l'exécution de divers scripts, par exemple, produire une image au format gif à partir d'un fichier créé par l'outil de dessin vectoriel Xfig, ou encore transformer un titre de section ou de chapitre en image en utilisant un des nombreux scripts de l'outil de traitement graphique GIMP.

La balise `<insérer>` permet d'insérer tout ou partie d'un fichier texte dont le nom est donné par l'attribut `fichier`. Deux autres attributs, `début` et `fin` permettent d'identifier deux chaînes de caractères, présentes dans le fichier texte et identifiant le début ou la fin de la partie à insérer.

### Dimension format

La dimension "format" permet de choisir le format dans lequel le document sera transformé: html, pdf, zip-html, ...

Au moment de la rédaction de cet article, plusieurs formats sont possibles: le format fo, le format pdf, le format html et le format zip. Ce dernier format permet la création d'un fichier zip contenant toutes les pages du document au format html, images comprises, permettant une exportation html facile d'un document pour un profil donné.

### Dimension filtre

Cette dimension est la plus originale. Elle permet de sélectionner un sous ensemble d'éléments du fichier initial suivant un ou plusieurs critères quelconques. Ces critères sont définis dans la déclaration d'un filtre apparaissant en début du document xml. On pourra, par exemple, autoriser l'accès à un élément particulier seulement à partir d'une machine du réseau d'un domaine donné, ou seulement aux utilisateurs identifiés sous un nom donné (l'identification se faisant de manière classique en début de session http, grâce à un mot de passe), ou bien pour tout accès après une certaine date.

### Serveur dynamique

L'intérêt de l'outil PubliTool est de pouvoir être utilisé également en mode serveur grâce à une servlet Java. Le principal avantage vient du fait qu'il n'est pas nécessaire de maintenir, en plus des documents xml, un site accessible via http.

### Chaîne de transformations XSL

PubliTool est dynamique dans le sens où les fichiers servis sont construits dynamiquement en appliquant une chaîne de transformations XSL. Les feuilles de styles utilisées peuvent être sélectionnées de plusieurs manières:

- elles peuvent être associées au format de sortie et au nom de l'élément racine du document xml servie. Par exemple, pour un document dont la racine a pour nom `article` et dont le format choisi est `html`, PubliTool utilisera de manière implicite la feuille de style ayant pour chemin d'accès `/XSL/html/article.xsl`.
- elles peuvent être désignées dans le document lui-même grâce à des instructions de traitement XML

Si plusieurs feuilles de styles sont référencées par des instructions de traitement XML, elles formeront une chaîne de transformations: le résultat de la transformation associée à la première instruction de traitement est utilisé comme source de la transformation associée à la deuxième instruction de traitement. Cette simple approche permet de définir des macros à la manière de latex. La version longue de cet article montre comment réaliser une image au format svg, en utilisant des feuilles de styles xsl, à la manière des macros latex.

### Filtres d'accès aux fichiers

Le principe du filtre s'applique également sur l'ensemble des fichiers ou répertoires accessibles par PubliTool. Il suffit d'ajouter dans un répertoire, un fichier de nom `"accèsRépertoire.xml"` contenant les filtres uti-

lisés pour les différents éléments du répertoire.

### Cache

La gestion dynamique n'interdit pas l'utilisation de cache pour améliorer les performances du système. PubliTool possède une fonction de cache limitant à une seule transformation, toute réponse à une requête conduisant à la même génération de document, qu'il soit interne à un document xml (comme une image générée à partir d'une balise latex) ou qu'il soit le document complet lui-même. Une réponse mémorisée dans le cache est alors renvoyée aussi rapidement que ne le ferait un serveur statique de document.

Pour chaque requête reçue, PubliTool construit une clé à partir d'un certain nombre d'informations. Par exemple, nous trouverons dans la construction d'une clé, les composantes suivantes: les dates de dernière modification de tous les fichiers liés au document, les valeurs de tous les filtres référencés dans le document, la partie paramètre (query) d'une requête http, une signature (crc) calculée sur le contenu d'un élément.

A chaque requête d'accès, PubliTool calcule une clé puis recherche si un fichier associé à cette clé a déjà été généré. Si c'est le cas, le fichier est servi, sinon il y a génération d'une nouvelle réponse dynamiquement.

### Conclusion

Nous avons présenté PubliTool, un serveur dynamique de documents multi-formats XML. Par rapport à des applications analogues, comme l'application Cocoon-Apache, PubliTool présente plusieurs avantages:

- il est léger (une dizaine de classes java pour la version de base),
- il utilise directement l'API version 1.4 du système de développement Java pour la partie XML,
- il intègre un accès protégé au niveau de l'élément interne d'un documents,
- il met en oeuvre une chaîne de transformations exprimée simplement,
- il permet de s'affranchir du sacro-saint copie-coller.

PubliTool est en fait la troisième version d'une application anciennement dénommée PolyDoc, utilisée depuis plusieurs années à l'INSA de Toulouse pour la création et la publication de documents pédagogiques. La première version de PolyDoc était basée sur le traitement d'un document à travers sa représentation DOM. La deuxième version était construite autour de l'application Cocoon 1 d'Apache et était basée autant sur la représentation DOM du document que sur la transformation de type XSL. Cette troisième version reporte la grande majorité du traitement sur la transformation XSL, et est entièrement compatible avec la version 1.4 du système de développement Java.

L'utilisation de PubliTool est libre (au sens GPL du terme). La distribution est téléchargeable à l'adresse <http://www.dgei.insa-tlse.fr/~castan/PubliTool.html>.