

Licence Pro - Technologies Web - Technologies XML - Programmation – Devoir Maison

Transformation et programmation XML

1.1 Objectifs

Il vous est demandé de réaliser la transformation d'un document XML en XHTML selon deux approches aux choix parmi celles ci-dessous. Le document XML est une bibliographie. Les choix sont

- à l'aide d'une feuille XSL que vous nommerez *BiblioToXhtml.xsl*
- à l'aide d'un parseur SAX (via l'API JAXP) dont vous nommerez la classe contenant le main *BiblioToXhtmlSax.java*
- à l'aide d'une manipulation DOM/XPath (via l'API JAXP) dont vous nommerez la classe contenant le main *BiblioToXhtmlDom.java*
- à l'aide d'une manipulation via un binding Xschema à l'aide d'une classe contenant le main que vous nommerez *BiblioToXhtmlJAXB.java* (vous fournirez aussi le XSchema réalisé et l'API générée correspondante).

1.2. Réalisation demandée

Le document XML à disposition contient une liste de conférences chacune contenant une liste de références de publications (articles). Pour la plupart des conférences, on trouvera son acronyme, son titre, ses dates, ses présidents, des statistiques et l'identifiant des meilleurs articles. Pour la plupart des publications on trouvera des auteurs, des affiliations, un titre, des mots clefs et un résumé.

Le document HTML à produire comptera deux divisions. Une première division jouera le rôle d'*index de conférences*, la seconde de *détail de chaque conférence*.

La division *indexConferences* listera les titres des conférences par ordre alphabétique et pour chacune indiquera les années (à extraire de la date de début) pour lesquelles il y a des références et ce dans l'ordre croissant de celles-ci. La mise en forme de l'index sera similaire à [1].

La division *detailConferences* contiendra une liste de divisions titrées par l'acronyme de la *conférence*.

La liste de divisions conférence sera ordonnée sur l'année (à extraire de la date de début) puis selon l'ordre alphabétique de l'acronyme de la conférence. Chaque division *conférence* contiendra une liste de *références* qui suivra l'ordre alphabétique du nom du premier auteur. La mise en forme des références sera similaire à [2].

Chaque division conférence s'accompagnera d'une en-tête avec les informations suivantes :

l'année (à extraire de la date de début), l'acronyme, les noms des présidents et le lieu de la conférence. Si il y a des statistiques, il faudra indiquer le nombre de soumissions et le nombre d'articles acceptés selon la catégorie d'articles sinon indiquer qu'aucune statistique n'est disponible.

Qu'il y ait ou non des statistiques, il faudra indiquer le nombre d'articles total référencés.

Si il y a un meilleur papier, il faudra l'indiquer sinon indiquer qu'aucune information n'est disponible à ce sujet.

Des années de l'index de chaque conférence (*indexConferences*) partiront des liens vers chaque titre des divisions *conférence*.

Il y aura aussi un lieu de l'information d'en-tête indiquant qu'il y a un meilleur papier vers l'emplacement de la description de celui-ci.

Pour aller plus loin (questions bonux)

* Ne pas faire une division détail contenant des divisions pour chaque conférence mais une page HTML dédiée pour chaque conférence (et par conséquent un lien pour y accéder à partir de l'index)

* Rajouter une division index des noms d'auteurs avec le nombre de publication de chaque auteur quelque soit la conférence

Références

[1] <https://aclweb.org/anthology/>

[2] <https://aclweb.org/anthology/D/D08/>

1.3 Livrables

Il vous est demandé une archive nommée *groupe.nom1.nom2.tgz* et contenant un répertoire *groupe.nom1.nom2* (avec *groupe* pouvant prendre les valeurs *g1* ou *g2*, et avec *nom1* et *nom2* respectant l'ordre alphabétique) avec

- un rapport compte rendu CR.txt (avec vos noms en en-tête) indiquant pour chaque approche (celles que vous avez utilisées et les deux autres) pourquoi vous l'avez retenue ou bien mise de côté. Vous vous appuyerez pour cela d'exemples de difficultés concrets que vous avez rencontrés et comment vos choix techniques vous ont permis de les résoudre. Indiquez aussi un état d'avancement des réalisations.
- les codes implémentant votre solution (une feuille XSLT, un code java implémentant l'approche SAX, un code java implémentant l'approche DOM, API générée et feuille XSchema pour l'approche JAXB)
- un fichier README.txt avec les commandes à taper dans un terminal expliquant comment compiler et exécuter votre code
- inclure le fichier exemple dans l'archive pour faciliter le test de vos codes.

1.4 Consignes

Ce devoir est à réaliser en binôme.

Envoyez un mail avec vos noms à votre chargé de TD pour déclarer votre binôme avant FIXME (à moins que celui-ci ait pris vos noms directement en classe). En cas de monôme veuillez vous signaler rapidement aussi par mail.

Les livrables seront à déposer sur Madoc au plus tard le FIXME.

Si besoin, vous avez le droit d'utiliser des outils ou technologies complémentaires à ceux présentés (dans ce cas, présentez votre utilisation dans votre rapport).