

Master 2 Humanit s num riques – Bases de donn es semi-structur es

TD 4 : XQuery – Requ tes XPath

J. Darmont – <https://eric.univ-lyon2.fr/jdarmont/>

Pr ambule

Soit un catalogue d'ouvrages stock  dans un document XML dont la structure hi rarchique est donn e ci-dessous. @ marque un attribut, * indique une multiplicit  « plusieurs » et ? une multiplicit  « z ro ou un »

```
catalog
.....book*
.....@id
.....author
.....title
.....genre
.....price
.....publish_date
.....description
.....onsale?
```

1. T l charger le document XML suivant et le stocker localement.

<https://eric.univ-lyon2.fr/jdarmont/docs/books.xml>

2. Lancer BaseX depuis le menu D marrer de Windows ou gr ce au paquetage JAR disponible   l'adresse suivante.

<https://basex.org/download/>

3. Cr er une nouvelle base de donn es uniquement avec le document books.xml (menu Database/New).

Exercice 1

Les requ tes XPath doivent  tre  crites dans l' diteur (fen tre en haut au centre, par d faut). Pour ex cuter, une requ te, utiliser les touches **Ctrl + Entr e** ou cliquer sur le triangle vert. Formuler les requ tes suivantes   l'aide d'expressions de chemin uniquement ! Pas de requ tes FLWOR pour l'instant.

1.  l ments *book* du document books.xml. Sp cifier le chemin complet des  l ments *book*.
2. M me question sans sp cifier le chemin complet des  l ments *book*.
3. Titres de tous les livres.
4. Identifiants de tous les livres.

5. Caractéristiques du 4^e livre.
6. Titres des 5 premiers livres.
7. Titres des livres dont le genre est « Computer ».
8. Livres de genre « Computer » et coûtant moins de 40 €.
9. Livres actuellement en solde (c'est-à-dire, qui possèdent un élément *onsale*).
10. Titres et descriptions des livres écrits par Eva Corets.
11. Prix moyen des livres.
12. Titres des livres publiés en 2001.
13. Livre le plus cher du catalogue.

Exercice 2

Télécharger l'extrait de données d'enchères réelles à l'adresse ci-dessous, créer dans BaseX une nouvelle base de données avec ce document et visualiser sa structure.

<https://eric.univ-lyon2.fr/jdarmont/docs/auctions.xml>

Formuler ensuite les requêtes suivantes avec XPath.

1. Nombre d'enchères (*listing*).
2. Nombre d'enchérisseurs (*bidder_name*).
3. Nombre d'enchérisseurs distincts.
4. Noms et évaluations (*seller_rating*) de tous les vendeurs (*seller_name*).
5. Évaluation la plus basse (ne pas tenir compte des nouveaux vendeurs).
6. Noms des vendeurs ayant l'évaluation la plus basse.
7. Objets avec un processeur (CPU) Celeron.
8. Nombre moyen d'enchères (*num_bids*).
9. Information à propos des objets sans élément mémoire (*memory*) spécifié.
10. Plus haute enchère (*current_bid*).