

# Examen de BDA, M1 2010

31 mai 2010

**Les documents sont permis. Le barème est donné à titre indicatif et peut être modifié. ECRIVEZ LE NOM DE VOTRE PARCOURS DANS LA PARTIE CACHEE DE LA COPIE, A COTE DE VOTRE NOM.**

## Exercice 1 : Modélisation Semi-Structurée, 7 points

On considère à nouveau la base de données qui gère les emprunts des livres d'une bibliothèque universitaire étudiée dans le DS. Mais, cette fois-ci, on veut une modélisation semi-structurée, et on précise certaines hypothèses.

La base contient des informations sur les étudiants, les livres, et les emprunts pour une période donnée, disons février 2009 (donc l'information sur la période n'apparaît pas dans la base). Un étudiant a un nom, un prénom, un numéro de carte d'étudiant, une adresse e-mail, parfois un numéro de téléphone (pas plus qu'un, si il en a un), et un ou plusieurs départements (un étudiant peut faire, en même temps, un L3 Math et un M1 Informatique, par exemple); des départements, on connaît juste les sigles (INF, BIO, MATH) qui agissent comme des identifiants. Un livre a un code, un titre, un ou plusieurs auteurs (dont on connaît juste le numéro de sécurité sociale -le NSS-, le nom et le prénom) et une maison d'édition (dont on connaît juste le nom). Dans la période en question, un étudiant peut emprunter au maximum un livre. *Ici, on supposera aussi que la bibliothèque a un seul exemplaire de chaque livre.* Donc un livre peut être emprunté par un étudiant au maximum, et son code l'identifie de façon unique. On supposera aussi que la base répertorie au moins un livre (donc au moins un auteur) et au moins un étudiant (donc au moins un département, aussi).

1. Donner un DTD adéquat pour cette base, en essayant de tenir compte au maximum des contraintes explicites ou implicites de la description de la base donnée ci-dessus.
2. Si des contraintes ne sont pas exprimées par votre DTD, on en fera la remarque, et on expliquera pourquoi elles ne sont pas exprimées.

## Exercice 2 : Requêtes Xquery, 10 points

On considère à nouveau la base de données mondial.xml vue en TD, dont le DTD est donné en annexe. Ecrire les requêtes Xquery suivantes :

1. Donner les noms des pays où la religion "Armenian Apostolic" est suivie par une partie de la population, et donner aussi les noms des langages qu'on y parle (mais pas l'information sur les *percentages* de ces langages).
2. Donner les noms des religions qui sont communes à la France et à la Georgie (*Georgia*).
3. Donner les noms des religions pratiquées en Georgie mais pas en France.
4. Donner les noms des pays dont la religion musulmane (Muslim) n'apparaît pas dans l'ensemble de ses religions.
5. Est-il vrai que la surface (*area*) de la France est plus que 200 fois celle du Luxembourg ?

6. Afficher les noms des premiers 4 pays (selon l'ordre dans le document) avec ce format :
  - 1 Nom\_du\_premier\_pays
  - 2 Nom\_du\_second\_pays
  - 3 Nom\_du\_troisième\_pays
  - 4 Nom\_du\_quatrième\_pays
7. Pour chaque pays, afficher le nom de sa capitale, selon le format de résultat indiqué ci-dessous :
  - \*\*
  - La capitale du pays nommé Albania est Tirane
  - \*\*
  - La capitale du pays nommé Greece est Athens

### Exercice 3 : Intégration de Données, 3 points

On veut gérer un comparateur de prix pour la vente de plantes de jardin sur internet. On a deux sources différentes, qui stockent leur données en XML. Les deux sources contiennent, respectivement :

Source 1 : des informations sur les plantes vendues par un fournisseur donné, disons Fournisseur1. Les seules informations de cette source portent sur Clé\_Fournis, Nom\_Fournis, Ville\_Fournis, Nom\_plante, Prix, où Clé\_Fournis, Nom\_Fournis et Ville\_Fournis renseignent sur Fournisseur1.

Source 2 : la même sorte d'informations que l'autre source, mais pour un autre fournisseur, disons Fournisseur2.

L'utilisateur interroge un comparateur de prix, qui doit fournir, pour chaque plante, le prix le plus bas parmi les deux fournisseurs. On fera l'hypothèse que les deux fournisseurs vendent exactement les mêmes plantes, mais à des prix éventuellement différents. Le schéma du comparateur est donc le schéma global ; Clé\_Fournis, Nom\_Plante, Prix\_Le\_Plus\_Bas, sont les seules sortes d'information du schéma global. Observer que Clé\_Fournis stocke l'identifiant du fournisseur qui vend la plante au prix le plus bas. Le langage du schéma global est toujours XML.

1. Ici, l'approche GAV est possible. L'approche LAV, est-elle possible aussi? Justifier votre réponse.
2. Ici, on ne considère que l'approche GAV.
  - (a) Donner deux DTD qui définissent les schémas respectifs des deux sources.
  - (b) Donnez une requête Xquery qui calcule les couples de valeurs de (Clé\_Fournis, Nom\_Plante, Prix\_Le\_Plus\_Bas) en interrogeant les deux sources.