

Câu 1: (3đ)

Hãy soạn một tài liệu DTD và một tài liệu XSD để tài liệu XML **books.xml** sau là hợp lệ:

```
<?xml version="1.0"?>
<catalog>
  <book id="bk101">
    <author>Gambardella, Matthew</author>
    <title>XML Developer's Guide</title>
    <price>44.95</price>
    <publish_date>2000-10-01</publish_date>
    <description>An in-depth look at creating applications with XML.</description>
  </book>
  <book id="bk102">
    <author>Ralls, Kim</author>
    <title>Midnight Rain</title>
    <genre>Fantasy</genre>
    <price>5.95</price>
    <publish_date>2000-12-16</publish_date>
  </book>
</catalog>
```

Câu 2: (3đ)

Cho file **books.xml** như câu 1:

- Hãy viết các biểu thức XQuery để trích rút ra các **title** của các phần tử **book** có **price** nhỏ hơn 30.00, có sắp xếp theo title
- Hãy viết một đoạn script theo mô hình đối tượng XML DOM hiển thị ra **title** của cuốn sách có **price** nhỏ hơn 30.00

Câu 3: (2đ)

Cho một tập các chuỗi $S = \{s_1, s_2, s_3\}$ với $s_1 = \text{"Loperamide"}$, $s_2 = \text{"Furosemide"}$ và $s_3 = \text{"Furosol"}$. Theo cách tính độ tương đồng Jaccard với 3-grams thì hai chuỗi nào gần nhau nhất? Trình bày cách tính này.

Câu 4: (2đ)

Hãy trình bày kiến trúc hệ tích hợp dữ liệu.

-----HẾT-----

Soạn đề thi

Duyệt đề thi