Two hours - on line

The exam will be taken on line.
This paper version is made available as a backup

QUESTION PAPER MUST NOT BE REMOVED FROM THE EXAM ROOM

UNIVERSITY OF MANCHESTER
SCHOOL OF COMPUTER SCIENCE

Semi-structured Data and the Web

Date: Monday 13th January 2014
Time: 09:45 - 11:45

Please answer ALL Questions provided.

The exam contains MULTIPLE CHOICE, TRUE/FALSE and SHORT ESSAY QUESTIONS.
Be sure to answer ALL Questions.

Please note that wrong answers on MULTIPLE CHOICE and TRUE/FALSE questions may be penalized (i.e. receive some small negative mark) so random guessing works against you.

This is a CLOSED book examination
The use of electronic calculators is NOT permitted
Questions 1-15 are restricted and cannot be published
16. Describe “well-formedness of an XML document” in one or two sentences.
   (2 marks)

17. Describe “the DOM tree or infoset of an XML document” in one or two sentences.
   (2 marks)

18. Describe “a validating XML parser” in one or two sentences.
    (2 marks)

19. Describe “a PSVI” in one or two sentences.
    (2 marks)

20. In some situations, a schema that corresponds to a single-type grammar can be more appropriate than one that corresponds to a grammar that is not single-type. Discuss, in 4-5 sentences, such a situation. Also mention why choosing such a single-type schema could cause problems.
    (5 marks)

21. Discuss, in 4-5 sentences, whether it is advantageous for XML to have many different schema languages.
    (5 marks)
22. In 4 - 5 sentences, describe four differences between Schematron and Relax NG as XML schema languages. (5 marks)

23. Write a DTD that describes the following XML document. Strive for a fairly tight, but sensible DTD. (It should be about 6 declarations long. Minor syntax errors and variations are ignored.)

```
<library owner='bob'>
    <item>
        <kind>book</kind>
        <details>
            J. K. Rowling: Harry Potter and the Goblet of Fire
        </details>
        <lent-to>
            Peter
        </lent-to>
    </item>
    <item>
        <kind>cd</kind>
        <details>P.J. Harvey: Let England Shake</details>
        <lent-to>nobody</lent-to>
    </item>
</library>
```
24. In 4 - 5 sentences, discuss the quality of the following two formats for a library, as exemplified in the following example documents. Also, suggest further improvements to the better format.

Format 1:
<library owner='bob'>
  <item>
    <kind>book</kind>
    <details>
      J. K. Rowling: Harry Potter and the Goblet of Fire
    </details>
    <lent-to>
      Peter
    </lent-to>
  </item>
  <item>
    <kind>cd</kind>
    <details>P.J. Harvey: Let England Shake</details>
    <lent-to>nobody</lent-to>
  </item>
</library>

Format 2:
<library owner='bob'>
  <friend name='Peter'/>
  <book>
    <author>
      J. K. Rowling
    </author>
    <title>
      Harry Potter and the Goblet of Fire
    </title>
    <lent to='Peter'/>
  </book>
  <cd>
    <artist>
      P.J. Harvey
    </artist>
    <title>
      Let England Shake
    </title>
    <lent/>
  </cd>
</library>

(6 marks)