CSE 8394
Assignment 1

Due: Feb 3. (2 weeks)

This assignment is intended to get you up to speed on XML technologies which serve as the basis for the semantic web technologies we will encounter. It is important that you have a solid foundation in XML technologies.

Part A. Work through the following tutorials at w3schools.com:

XML: http://www.w3schools.com/xml/

DTDs: http://www.w3schools.com/dtd/default.asp

Part B.

Define at least two XML vocabularies for data about some domain (one that is not the domain you may be thinking about for your final project). Some suggestions:

- model parts of your workplace/school - for example, if at a university, design a vocabulary about courses, teaching staff, rooms, publications. [example: faculty.xml, student.xml, courses.xml]
- model one of your hobbies - design a vocabulary for exchanging information with others who share your interest
- do you like cooking and/or wine? - design a vocabulary for exchanging information with others who share your interest [example: wine.xml, vineyard.xml,]
- do you invest? design a vocabulary about stocks, bonds, investment options, risk, etc.

Key questions to ask yourself:

- What data will I capture? What kinds of searches do I want to make?
- How will it be structured? You should have both elements and attributes in your XML.
- You should come up with at least two different xml files.

When you define your basic XML:

- create DTD(s) that defines your XML
- define a namespace and use it in your XML
- use ID attributes to define keys for your data
- use IDREFS to refer to entries with IDs
- define and use at least one entity and entity reference
- validate your XML against your DTD(s)

Part C.

- Select one or more data items that involve enumeration and/or numeric constraints. For example, an XML element called <school>school-name</school> may be defined as #PCDATA (text) but be
limited to one of several school names (LyleEngineering, MeadowsArts, Dedmon, Business, Law, Theology). Or, an element called `<grade>`95</grade>` may be constrained to an integer between 0 and 100.

- For enumerations and numeric values for which constraints apply, write XML Schema representations for these values.