

1. [10 points] Choose a document or piece of text that you're interested in from your workplace. This document can be structured or unstructured. You may choose a text of any sort and any length, as long as it is long enough to meet the following encoding criteria. Be sure to include the text as a separate file in your upload (or if it is online, you may provide a link to it)
2. [20 points] Make an XML DTD for that text. Your DTD should specify the following:
  - At least 10 different elements
  - At least 5 different attributes (you may have multiple attributes per element; not every element requires attributes). For at least 2 attributes, a controlled list of values the attribute may take.
  - Indicate, of course, whether each element and attribute is required, optional, an either/or, repeated, etc.
  - Indicate how elements may nest.
3. [10 points] Write prose documentation for each element, attribute, and attribute value.
4. [10 points] Mark up the text you have chosen according to the DTD you designed. Make sure you include either a `<!DOCTYPE>` reference to an external DTD file, or an internal DTD within the `<!DOCTYPE>` element in the XML document. Be sure to validate your document (see below) as validation is one of the areas you will be assessed on. Be sure to validate your document (see below) as validation is one of the areas you will be assessed on.
5. [25 points] Write a narrative about this process, answering the following reflection questions:
  - How did you decide to represent the data in the way that you did? Why did you choose the elements and attributes that you did?
  - What were the hardest decisions you had to make in this design process?
  - How does your DTD design support data independence?
  - How may your DTD design support the overarching goals of data curation (revisit objectives and activities of Week 1)?
  - What are the pros and cons of your DTD design?
6. Each student will be required to grade the submissions of 5 of their peers. Submissions will be graded based on the following criteria. Write a constructive and professional review and post to the course forum replying to the individual's submission.
  - a. Is everything represented?
  - b. Is it clearly written? Is the scheme and data clearly presented?
  - c. What are the pros and cons of this representation?
  - d. How could it be done differently? How could it be improved?

## Assignment 2 – XML Schema Design Exercise

- RAJA SHEKAR BOLLAM (bollamr2)

**For Assignment 2, I choose to create a DTD document for my Resume. I would like to create XML Elements for each sections and details from my resume document.**

Below are the documents that are submitted for the assignment purpose.

- DTD Files
- DTD + XML File Integrated
- Sample Document (Resume)
- Generated XML output

### Explanation

DocumentType **resume** comprises of sections like personal, introduction, education, profession experience, technicalskills and others. Below is the explanation for each sections of the DTD.

#### Personal

This section has 3 attributes Name, Phone number and Email ID grouped together as personalInfo. I have marked all these attributes values a REQUIRED. It was intentional to mark each attribute as Elements to understand the shortcomings of it. Also I have empty initialized all the attributes. Note that atleast one section of this element should be present in the document.

#### Introduction

This section is an introductory note of the person. It is more of a summary of the resumes owners. This section is categorized as ANY, meaning it contain any combination of parsable data.

#### Education

This section is intentionally made as plain an simple. It has at least **ONE** element institute that constitutes of instinstitute Name, Type, Year of Starting, Year of Ending and Major. Each item is an parsable element in itself.

#### ProfessionalExperience

This section details the experience in related to profession. This section is grouped as Employer, Designation, Project, YearStart, YearEnd and JobDescription. YearEnd is labeled as an Optional field as an individual could be still working in the company. All the other fields are marked as Elements.

#### TechnicalSkills

This section we group the technical capabilities of the candidate. In this section we use the proper way of using a attributes unlike from the section “*personal*”. *Domain, Languages, skills, protocols, databases and Operating System* information is grouped together in this section. Notice that we have marked “databases” and “OS” as “*IMPLIED*” attribute type. To explain about default initialization we have default values for skills.

#### Others

This section is to have any miscellaneous information that needs to be presented. Notice that this section is marked with (\*). This is an optional element.