I have chosen an intent letter for job application as the document to implement this activity.

- At first the document has been divided into 3 main elements which are the head, body and footer.

- The head of the letter consists of the following elements:

1) SendDate which carries information on when the letter was sent to the employer

2) Recipient which carries four sub-elements which are:

- Name - contains three attributes: Title(Optional), FirstName(Required) and LastName(Required)

- Position has information about the job position of the Recipient

- Company of the Recipient

- Address of the Company contains 4 elements which are Street, City, State and Zip

-The body of the letter contains the actual content of the letter and has two sub-elements:

1) Salutation has two definitive attributes; Greet and Name to letter is being referred to.Both of which are compulsory to mention.

2) Para has content related to the employee seeking the job. There can be one or more Paras

-The footer contains the closing part where the job seeker gives his information, and it has 3 elements:

1) Closing has a required attribute Valedictory with a value assigned

2) SenderName has two required attributes FirstName and LastName

3) SenderAddress further contains 4 elements SenderStreet, SenderCity, SenderState, SenderZip

We see that since there is name and address repeated for the recipient and the sender. To avoid elements from overlapping, the Sender prefix is added to the latter.

* 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?

The letter has been subdivided as 3 components -head, body and footer which is a standard format for most letters. Within each elements there are other related elements to breakdown the complexity of arrangement. There are 23 elements and 8 attributes in total. The elements are the main components into which the letter is divided and the attributes support the elements by providing more information about the elements. The head of the letter contains information about the recipient. The body contains the intent of the letter and the footer contains the sender’s information. We see that since there is name and address repeated for the recipient and the sender. To avoid elements from overlapping, the Sender prefix is added to the latter. The attributes are named in a way that they give more information about the elements that they are present within. For example, the Name element in the head has three attributes which take on controlled list of values. They are the Title, FirstName and LastName. The Title can be optional, however the First and Last Name are required. The elements are defined as parsed character data, however their attributes are character data.

* What were the hardest decisions you had to make in this design process?

The hardest decision was of course to decide whether to consider an object as element or an attribute by analyzing what best suits it and the avoid too many unnecessary attributes.

* How does your DTD design support data independence?

The DTD supports abstraction by explicitly identifying the recurring objects and thus abstracts away from the details of storage and processing.

Indirection- We only need to specify the logical components of the data in form of attributes and elements and the system exploits the mapping by indirection to give the effects.

* How may your DTD design support the overarching goals of data curation (revisit objectives and activities of Week 1)?

The DTD discussed above helps achieve data analysis in a much effective manner due to its standard format for similar documents. Also, the format can be reused over time with more incoming data.