**Question 4:**

I have chosen the benefits claim submission form as the document for my DTD.

The main element for this DTD is BENEFITDATA. This has different elements in it which are categorized into EMPLOYERDATA, EMPLOYEEDATA, PATIENTDATA, CLAIMDETAILS, MEMBERDATA, MEDICALAUTH, PAYMENTAUTH

**EMPLOYERDATA:** This is the first element section and represents the employer data. It has two additional elements in it EMPLOYERNAME and GROUPNO. Both these are required columns

**EMPLOYEEDATA:** This is the second element section and represents the employee data. It has additional elements EMPLOYEENAME, EMPLOYEEADDRESS and EMPLOYEEDAYPHONE. EMPLOYEEADDRESS has additional nested elements EMPADDRESS1, EMPADDRESS2, EMPCITY, EMPSTATE and EMPZIPCODE. EMPLOYEEDATA has attributes like EMPIDNO, EMPLOYEEDOB and RETIREMENTSTATUS. EMPIDNO and EMPLOYEEDOB are required attributes. RETIREMENTSTATUS has a enumerated list with “Active and Retired” with a default value of “Active”.

**PATIENTDATA:** This is the third element section and represents the patient data. It has additional elements such as PATIENTNAME, PATIENTADDRESS, PATIENTEMPLOYER, PATIENTEMPADD and OTHER\_INSURANCE. PATIENTNAME is the only required element. All other elements are implied. PATIENTADDRESS has additional nested elements PADDRESS1, PADDRESS2, PCITY, PSTATE and PZIPCODE which are also optional fields. PATIENTEMPADD also has additional nested elements such as PEADDRESS1, PEADDRESS2, PECITY, PESTATE and PEZIPCODE. These are optional fields.

PATIENTDATA also has attributes PATIENTIDNO, PATIENTDOB, PATIENTRELATION, PATIENTGENDER, PATIENTMARRIAGE, PATIENTEMPLOYED and OTHER\_INSCHECK. PATIENTIDNO and PATIENTDOB are required attributes. PATIENTRELATION has an enumerated list with “Self, Spouse, Child, Other” as values and "Self" as the default value. PATIENTGENDER has an enumerated list with “Male, Female”. PATIENTMARRIAGE has an enumerated list “Married, Single” with a default value "Single". PATIENTEMPLOYED has an enumerated list “No, Yes” with a default list "No". OTHER\_INSCHECK has an enumerated list “No, Yes” with default value "No".

**CLAIMDETAILS:** This is the fourth element section and represents the claims data. This has additional elements CLAIM\_ACCDATE and CLAIM\_ACCTIME. These field are optional. CLAIMDETAILS also has additional attributes CLAIMACCIDENT and CLAIMEMPLOYMENT. CLAIMACCIDENT has a list of values “No, Yes” with a default value "No". CLAIMEMPLOYMENT has a list of values “No, Yes” with default value "No".

**MEMBERDATA:** This is the fifth element section and represents the member data. This has an additional element MEMBERNAME and attributes MEMBERIDNO and MEMBERDOB which are REQUIRED attributes.

**MEDICALAUTH**: This is the sixth element section and represents the medical authorization for release of data. It has two additional elements MSIGNATURE and MDATE. This also has an attribute MDEFAULTMESSAGE with a default fixed message "To all providers of health care providers. You hear by authorize to provide medical data for claim processing".

**PAYMENTAUTH**: This is the last section and represents the payment authorization for the payment for the claim. It has two additional elements PSIGNATURE and PDATE. This also has an attribute PDEFAULTMESSAGE with a default fixed message "I authorize payment of medical benefits to physician or supplier of service".

**Question 6:**

**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 data I have used is the medical benefits claim form. I have grouped the data in the form into different groups such as employer data, employee data, patient data, claims details, member data, medical authorization and payment authorization. Grouping the data into different categories helps us to store the data easily in the database. Since its difficult to change the attributes, I have chosen columns that need not change easily as attributes. I have also chosen he ID and DOB column as attributes since these uniquely identify the rows. I have chosen the other columns as elements.

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

Since there is no definite definition to choose attributes and elements deciding which column would be an attribute or element was difficult. I was also debating if the date of birth column needed nesting. I then decided to have the data of birth column as an attribute.

**How does your DTD design support data independence?**

I have created the DTD by grouping them into different categories. This helps us to store data easily. Also, each data group can change independently without affecting the other.

**How may your DTD design support the overarching goals of data curation?**

The main goals of data curation are data maintenance and reliability for reuse and analysis. Since the data is categorized, data maintenance is easy. This also helps in reusing the data. Having certain columns as required helps with performing analysis.