**ENCODING INFORMATION IN I2b2 SCHEMA**

**Assumptions:**

1. Each CONCEPT\_CD in a CONCEPT\_DIMENSION table is unique.
2. Ontomapper maps each CONCEPT\_CD to a Common Data Element (CDE) from caDSR
3. There are multiple caDSR services present in a distributed environment where I2B2 operates.

**What we need to do:**

We would need to store the following information (attributes) in order to build the domain model from the CaDSR:

-Service\_URL

-Context\_Name

-Project\_Id

-Project\_Name

-Project\_Version

-CDE\_Id

-CDE\_version

**Reasons for selecting each of these:**

Service\_URL holds the URL for the caDSR Service pointing to any local or remote instance and uniquely identifies the Context.

Context\_Name owns the various projects within the Classification Scheme and uniquely identifies the Context to which the CDE points to.

Project\_Id stores the Public Id for the project to which a CDE belongs.

Project\_Name stores the short name of the Project and may be useful in identifying the project.

Project\_Version stores the information about the exact version of the project to which the CDE belongs, as there could be multiple versions of the same project existing within the caDSR.

CDE\_Id stores the public Id of the Common Data Element and is necessary for identifying it uniquely.

CDE\_version stores the exact version of the CDE within the Project.

The new schema could look something as below:

1. Addition of 2 new tables: ENCODING SERVICE and ENCODING\_PROJECT
2. Addition of 3 columns (CDE\_PUBLIC\_ID, CDE\_VERSION, ENCODING\_SERVICE\_ID) in ENCODING\_DIMENSION table

