

Init6pm23 - Brain Tumor in silico study - Pathology and Radiology data models

Pre Interview:

Item	Information/Response
Date:	12/18/2009
Requirement # unique id <SemCon Ops Initiative>. <analysts initials><requirement number> e.g. Init1dbw1 (eventually linked to Use Cases)	Init6pm23
Originator/Customer's Name:	Joel Saltz - forum post
Originator/Customer's Company:	Emory University
Summary of requirement pre-interview, by Reviewer:	A Brain Tumor in silico use case consists of determining genetic, gene expression and outcome correlates of high resolution nuclear morphometry in the diffuse gliomas and their relation to MR features using Rembarth and TCGA datasets. This involves integrative analysis involving Pathology, Radiology and molecular data. The following semantic infrastructure requirements fall out of this use case: • Specific scientific data elements will be shared amongst collaborators, requiring the need for a way to semantically describe the data. However, through the course of the study, new data elements will be added and some data elements may change. Therefore, there is a need for an agile modeling approach that does not require significant effort to modify the information model and register the semantic metadata. This will greatly facilitate the job of the Information Modeler. • Data elements are generated using specific algorithms. There needs to be a way to model the features of the algorithm itself and tie it back to the original data. One of the features of the algorithm could be the code of the algorithm itself. It would be ideal if this type of model could be generalized for use in the caBIG analytical community so that Software Engineers can reuse it and Cancer Researchers have access to this data. • The scientific data elements themselves will describe not only the statistical results, but the scientific methodology used on the specimens. In some sense, this combines the study design with the results. • The ultimate result of the scientific use case is to design a classifier that predicts outcome. It will be necessary to describe the classifier, as well as have that description link back to the data that is used to generate it. In other words, the provenance of the data must be captured so that Cancer Researchers have access to this information. A common model would facilitate development of systems by Software Engineers.
Recommended Next Step Enter one: Follow-up interview, Observe, Use Case Template (text), Use Case Model (formalized/UML diagram), Group Discussion, Prototype, Waiting Room	1. Use Case Template 2. Followup Interview