A Unique Opportunity in Biological Information Object Standards

C.F. Dewey, Jr.[1,2] and **Shixin Zhang**[1]

[1]Department of Mechanical Engineering
[2]Division of Biological Engineering

Massachusetts Institute of Technology, Cambridge MA USA

Introduction

Over the past several years, the explosive growth of biological data generated by new high-throughput instruments has literally begun to drown the biological community. There is no established infrastructure to deal with these data in a consistent and successful fashion. This paper discusses the opportunity to develop a new informatics platform to handle a large subsection of the experimental protocols that currently exist. A consistent data definition strategy is demonstrated that handles gel electrophoresis, microarrays, fluorescence activated cell sorting, mass spectrometry, and microscopy within a single coherent set of information object definitions. Other experimental methods can be added with relative ease because the object model used to describe the data is easily extended.

The next step in the development of this platform is to enable simple access and use over the World Wide Web. The two keys to this deployment are: (a) establishing consistent ontologies accepted by the biological community and made available from repositories using OWL technology; and (b) using the rich descriptive capabilities of RDF to exchange data between repositories and users. These two technologies will enable broad data sharing and interoperability within the biological community.

Methods

Several important experimental techniques in contemporary biology have been used to create a single composite schema. The results bear a striking relationship to the DICOM standard of 1993 that provides information object definitions of all of the major medical imaging modalities (MR, CT, US, XA, NM, VL, CR, and Waveforms). The *de novae* information object definition we developed for gel electrophoresis turned out to be very similar to the existing MAGE-OM information model for microarrays. Further investigation revealed that similar object definitions characterized other experimental biology methods as well. These were generalized and a full object-relational data schema was developed. The appended references cite a number of the proposed standards that were used to develop the object model.

Results
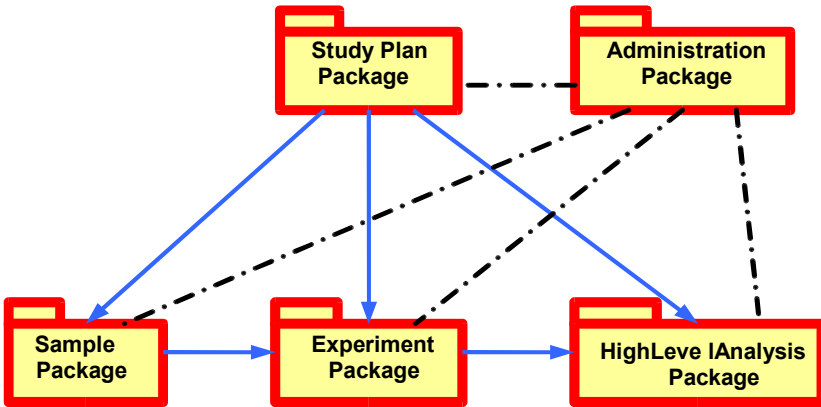
A ▶B Dependency. The changes of A can cause changes in B.
-.-.-.- Reference