Michel presented the dataset description model for unversioned -> versioned -> formatted -> data

- Overview: https://docs.google.com/drawings/d/1e6qsxPkc-qKecVTJGJePE1Nuy2sD8Puu-FsEtUtGC-o/edit?disco=AAAAAFWHcnw
- Detailed: https://docs.google.com/file/d/0B4y0zfdRviKsS1l2NEttN3pfc1k/edit?usp=sharing

Only an RDF data item would be able to point back to its description due to the allowances of the data model.

Discussion on whether to repeat metadata at all levels. Each description would be complete: simplifies query; redundancy in metadata. Flexibility of the URIs for each level means you have several entry points. Could have contradictions in the metadata at the different levels. Idea would be to limit the metadata in the unversioned part to minimal data that would not change over time.

Abstract data format, e.g. triplestore. One option would be to model it as a separate versioned/formatted dataset description. However there is an underlying formatted dataset that has been loaded into the the underlying datastore and this is what would be described. For a relational database accessible through D2R the description would be a SQL versioned dataset with an accessibility protocol of SPARQL. Service points to the versioned/formatted dataset that they expose.

Different syntaxes provide different views of the data. In RDF you can capture the relationship between the data and the metadata. This is not generally possible in other syntaxes.

Sources: points to the exact file that was used so that bugs can be tracked

MIRIAM is a catalog that described datasets. Catalog was added.

Format types would be captured with URIs.
EDAM are amenable to extending to cover the file type that we require, Nick has been in contact with Jon Ison. http://edamontology.org/page

ToDo: Go through different scenarios, e.g. Bio2RDF, Open PHACTS, MIRIAM and see how these look in the model (not using any particular vocabularies). By hand generate a full description for a dataset.

ToDo: Revisit properties in the spreadsheet to ensure that they are all still required.

Do send questions and comments to the list!