

# The CellML Metadata Specification

## SPMB Caltech June 1,2 – Discussion Proposal

Warren Hedley, The Bioengineering Research Group, The University of Auckland, April 23, 2001

### Current Status of CellML Metadata

To speed up the development of CellML, the scope of the CellML project was split in two in the second half of 2000. The first half of the project is the CellML Specification, a first draft of which was released at the start of March 2001, concerns itself with the definition of model structure and mathematics.

The second half of the project is the CellML Metadata Specification, a draft of which the CellML development team is currently working towards. The CellML metadata specification will describe an XML- and RDF-based syntax for the definition of the following types of information, which can be embedded within CellML documents.

- Human-readable names for CellML objects.
- Model authoring information, including
  - structured name information,
  - structured contact details information,
  - creation and modification date information,
  - publisher information,
  - copyright information.
- Annotations. These can be of type *comment*, *limitation*, *modification*, and *validation*, and may include
  - XML-based documentation formats (such as XHTML),
  - non-XML-based documentation formats (such as plain text or LaTeX),
  - structured annotation authoring information,
  - structured annotation date information.
- Relevant biological information, including
  - the species and sex with which a CellML object is associated,
  - the biological entity (an identifier for a protein or cell type).
- Mathematical metadata (a GAMS classification can be associated with a single equation, component or an entire model).
- Structured citation information.

### Discussion Proposal

We think the SPMB meeting at Caltech would be a great opportunity to promote the use of RDF for metadata languages throughout the field, to increase the chances of the metadata standards created in different projects being interoperable.

We would also like to present the data model and syntax defined in the CellML Metadata Specification to the SBW community. We would specifically like to encourage discussion on the minimal metadata data model, so that we can ensure that facilities for defining this information will be described in the final version of the specification.

We also hope to obtain feedback on the chosen syntax for metadata in CellML, particularly in areas where we have made use of existing standards and other languages. Currently these external

syntaxes include Dublin Core, V-Card and our own serialization of the OMG's Bibliographic Query Service (BQS).