

The CellML Metadata 1.0 Specification

Research Summary - 18 May 2001

2 General Metadata

The following subsections demonstrate the application of the recommendations in the previous three sections to all non-reference and non-person metadata. Note that the names of the mathematical problem type metadata elements have changed slightly, to remove a potential source of confusion. This is discussed in Section 2.11.

The discussion presented here assumes that the reader is familiar with the use of Dublin Core elements and Dublin Core qualifiers. These are discussed in the 26 March 2001 meeting minutes.

2.1 Alternative Names

Alternative name metadata provides human-readable names for CellML elements. One of these names can be considered to be the preferred name, equivalent to the old “displayname” concept in CellML99. This preferred name could be used by software whenever it needs to display a human-readable name. The use of this metadata allows us to limit the values of **name** attributes on CellML elements to enable efficient code generation, without worrying about whether or not the name will be sufficiently meaningful to human readers.

Alternative name metadata is defined with the Dublin Core title element, **<dc:title>**.

One element may have multiple alternative names. Only one should be considered the preferred human-readable name. The preferred name should be stored in an unqualified **<dc:title>** element. Additional names should be stored in **<dc:title>** elements that are qualified by setting the title type (**<dcq:title>**) to “alternative”.

Figure 6 shows the definition of alternative name metadata. The element referenced by ["#cellml_element_id"](#) is given two human-readable names. The preferred one is “EGF-EGFR complex”.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:dcq="http://purl.org/dc/qualifiers/1.0/">

  <rdf:Description about="#cellml_element_id">
    <dc:title>EGF-EGFR complex</dc:title>
    <dc:title rdf:parseType="Resource">
      <dcq:titleType>alternative</dcq:titleType>
      <rdf:value>
        epidermal growth factor-epidermal growth factor receptor complex
      </rdf:value>
    </dc:title>
  </rdf:Description>
</rdf:RDF>
```

FIGURE 6: Recommended definition of alternative name metadata.

2.2 Model Builder

Model builder metadata stores information about the person or persons who coded the model into CellML. A given element can have multiple model builders, which may need to be considered as individuals or as members of a group. If they are members of a group, the group may or may not need to be ordered.

Model builder metadata is defined using the Dublin Core creator element, **<dc:creator>**. Repeating this element for a given CellML element indicates that the people listed worked independently on the model. This definition is shown in Figure 7. Listing multiple people in the **<dc:creator>** element using an **<rdf:Bag>** container indicates that the group of people worked together on the model, and that they are all considered equal contributors. This definition is shown in Figure 8. Listing multiple people in the **<dc:creator>** element using an **<rdf:Seq>** container indicates that the group of people worked together on the model, and that their contributions are ordered (the first member of the list is first author, the second member is second author, and so on). This definition is shown in Figure 9. The CellML metadata specification will not include the use of an **<rdf:Alt>** container with this type of metadata. Metadata authors are free to use this container (as long as they produce valid RDF). However, CellML metadata compliant software is not required to be able to consistently interpret the meaning of an **<rdf:Alt>** container in this context. Note that in all of the examples shown here, the basic vCard “name” construct is used to store the name of the model builder. This and other vCard constructs will be discussed in a later document.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#">

  <rdf:Description about="#cellml_element_id">
    <dc:creator rdf:parseType="Resource">
      <vCard:N rdf:parseType="Resource">
        <vCard:Family>Flinstone</vCard:Family>
        <vCard:Given>Fred</vCard:Given>
      </vCard:N>
    </dc:creator>
    <dc:creator rdf:parseType="Resource">
      <vCard:N rdf:parseType="Resource">
        <vCard:Family>Brown</vCard:Family>
        <vCard:Given>Charlie</vCard:Given>
      </vCard:N>
    </dc:creator>
    <dc:creator rdf:parseType="Resource">
      <vCard:N rdf:parseType="Resource">
        <vCard:Family>Doo</vCard:Family>
        <vCard:Given>Scooby</vCard:Given>
      </vCard:N>
    </dc:creator>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 7: Recommended definition of model builder metadata in which multiple people worked independently on the model.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#">

  <rdf:Description about="#cellml_element_id">
    <dc:creator>
      <rdf:Bag>
        <rdf:li rdf:parseType="Resource">
          <vCard:N rdf:parseType="Resource">
            <vCard:Family>Flinstone</vCard:Family>
            <vCard:Given>Fred</vCard:Given>
          </vCard:N>
        </rdf:li>
        <rdf:li rdf:parseType="Resource">
          <vCard:N rdf:parseType="Resource">
            <vCard:Family>Brown</vCard:Family>
            <vCard:Given>Charlie</vCard:Given>
          </vCard:N>
        </rdf:li>
        <rdf:li rdf:parseType="Resource">
          <vCard:N rdf:parseType="Resource">
            <vCard:Family>Doo</vCard:Family>
            <vCard:Given>Scooby</vCard:Given>
          </vCard:N>
        </rdf:li>
      </rdf:Bag>
    </dc:creator>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 8: Recommended definition of model builder metadata in which multiple people worked together on the model, and all are considered equal contributors.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#">

  <rdf:Description about="#cellml_element_id">
    <dc:creator>
      <rdf:Seq>
        <rdf:li rdf:parseType="Resource">
          <vCard:N rdf:parseType="Resource">
            <vCard:Family>Flinstone</vCard:Family>
            <vCard:Given>Fred</vCard:Given>
          </vCard:N>
        </rdf:li>
        <rdf:li rdf:parseType="Resource">
          <vCard:N rdf:parseType="Resource">
            <vCard:Family>Brown</vCard:Family>
            <vCard:Given>Charlie</vCard:Given>
          </vCard:N>
        </rdf:li>
        <rdf:li rdf:parseType="Resource">
          <vCard:N rdf:parseType="Resource">
            <vCard:Family>Doo</vCard:Family>
            <vCard:Given>Scooby</vCard:Given>
          </vCard:N>
        </rdf:li>
      </rdf:Seq>
    </dc:creator>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 9: Recommended definition of model builder metadata in which multiple people worked together on the model, but all are not considered equal contributors. In this example, Fred Flinstone is the first author, Charlie Brown is the second author, and Scooby Doo is the third author.

2.3 Species

Species metadata refers to the biological species (such as human, dog, pig, etc.) for which an element is relevant. A given CellML element may be relevant for multiple species. It may also be relevant for an entire class of species, such as all mammals.

Species metadata is defined with a CellML-specific metadata element, **<cmeta:species>**, as shown in Figure 10. The content of this metadata must be a valid scientific name for a species or group of species. Notwithstanding recent arguments among taxonomists about the impact of genomic data on species classifications, scientific names are considered to be sufficiently standard to obviate the need to use a formal controlled vocabulary. Constructing such a vocabulary would be a daunting task! However, the CellML metadata specification will refer to the [NCBI's Taxonomy Browser](http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html)¹ as a good resource for scientific names. If a modeller needs to refer to a discontinuous group of species (i.e., one that cannot be specified by a single scientific name) he/she can include multiple **<cmeta:species>** elements. This was chosen over the use of RDF containers because it is simpler, and there was no need to differentiate between different possible meanings of multiple values for the species metadata. Multiple values for the species metadata will always mean that the CellML element is relevant for any one of the species listed. Relevance to all of the species as a group would imply some sort of population dynamics model, which is outside of the scope of CellML.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:cmeta="http://www.cellml.org/metadata/1.0#">

  <rdf:Description about="#cellml_element_id">
    <cmeta:species>Mammalia</cmeta:species>
    <cmeta:species>Xenopus laevis</cmeta:species>
  </rdf:Description>
</rdf:RDF>
```

FIGURE 10: Recommended definition of species metadata. The element referenced by "#cellml_element_id" is relevant for all mammals and the African clawed frog, *Xenopus laevis*.

2.4 Sex

Sex metadata refers to the sex for which a CellML element is relevant. A given element may be relevant for more than one sex.

Sex metadata is defined with the CellML-specific element, **<cmeta:sex>**, as shown in Figure 11. The valid content of this element must be chosen from the following controlled vocabulary:

- male
- female
- hermaphrodite
- undefined (the element is explicitly specified not to have a defined relevance to any particular sex).

¹<http://www.ncbi.nlm.nih.gov/Taxonomy/taxonomyhome.html>

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:cmeta="http://www.cellml.org/metadata/1.0#">

  <rdf:Description about="#cellml_element_id">
    <cmeta:sex>male</cmeta:sex>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 11: An example of the use of sex metadata

2.5 Creation Date

The creation date is the date upon which the model or model part was coded into CellML. A given CellML element can have only one creation date.

Creation date metadata is defined using the fully-qualified form of the Dublin Core date element, **<dc:date>**. The fact that the date is a creation date is indicated by setting the date type qualifier (**<dcq:dateType>** to “created”. The encoding scheme for the date is named in the date scheme qualifier (**<dcq:dateScheme>**). The allowed values of the encoding scheme qualifier are a controlled vocabulary from the Dublin Core (see [the Dublin Core Qualifiers document](http://dublincore.org/documents/dcmes-qualifiers/)².) The definition of creation date metadata is demonstrated in Figure 12.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:dcq="http://purl.org/dc/qualifiers/1.0/">

  <rdf:Description about="#cellml_element_id">
    <dc:date rdf:parseType="Resource">
      <dcq:dateScheme>W3C-DTF</dcq:dateScheme>
      <dcq:dateType>created</dcq:dateType>
      <rdf:value>2000-10-05</rdf:value>
    </dc:date>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 12: Recommended definition of the creation date metadata.

2.6 Last Modified Date

The last modified date is the date upon which the content of a CellML element was last changed. A given CellML element can have only one last modified date.

²<http://dublincore.org/documents/dcmes-qualifiers/>

The last modified date metadata is defined with the fully-qualified Dublin Core date element. Its definition is exactly the same as that of the creation date metadata, except that the value of the date type qualifier (**<dcq:dateType>**) is “modified”. The definition of creation date metadata is demonstrated in Figure 13.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:dcq="http://purl.org/dc/qualifiers/1.0/">

  <rdf:Description about="#cellml_element_id">
    <dc:date rdf:parseType="Resource">
      <dcq:dateScheme>W3C-DTF</dcq:dateScheme>
      <dcq:dateType>modified</dcq:dateType>
      <rdf:value>2000-10-05</rdf:value>
    </dc:date>
  </rdf:Description>
</rdf:RDF>
```

FIGURE 13: Recommended definition of the last modified date metadata.

2.7 Annotations

There are four types of annotations that will be recognized by the CellML metadata specification. Model authors are free to create additional types. However, CellML metadata compliant software will not be required to recognize any annotation types except for the following four:

- **comment:** free-form comment of the person who coded the model into CellML.
- **limitation:** brief description of the limitations/scope of the content of the CellML element.
- **modification:** description of a change made to the content of the CellML element.
- **validation:** description of the level of validation of the content of the CellML element. This may be a code. Note that validation codes are unlikely to be interoperable.

Each annotation also has creator and creation date metadata that refers to it.

Annotation metadata is defined using a CellML-specific element, **<cmeta:annotation>**. This element is qualified to include a type (**<cmeta:annotation_type>**) that indicates which type of annotation is included. The content of the **<cmeta:annotation_type>** element is a vocabulary controlled by the CellML RDF specification, with four valid values: comment, limitation, modification, validation. If a model author wishes to use a different value, he/she must place the **<annotation_type>** element in an application-specific namespace.

The author metadata associated with an annotation is defined exactly as the model builder metadata (Section 2.2), and creation date metadata associated with an annotation is defined exactly as the general creation date metadata (Section 2.5).

Figure 14 demonstrates the definition of comment and limitation annotations. Figure 15 demonstrates the definition of modification and validation annotations.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:cmeta="http://www.cellml.org/metadata/1.0#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:dcq="http://purl.org/dc/qualifiers/1.0/"
  xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#">

  <rdf:Description about="#cellml_element_id">
    <cmeta:annotation rdf:parseType="Resource">
      <cmeta:annotation_type>comment</cmeta:annotation_type>
      <rdf:value>This model does not include the data of Jones, et al.
        about the corresponding pathway in canine.</rdf:value>
      <dc:creator rdf:parseType="Resource">
        <vCard:N rdf:parseType="Resource">
          <vCard:Family>PowerPuff</vCard:Family>
          <vCard:Given>Bubbles</vCard:Given>
        </vCard:N>
      </dc:creator>
      <dc:date rdf:parseType="Resource">
        <dcq:dateScheme>W3C-DTF</dcq:dateScheme>
        <dcq:dateType>created</dcq:dateType>
        <rdf:value>2001-04-01</rdf:value>
      </dc:date>
    </cmeta:annotation>
    <cmeta:annotation rdf:parseType="Resource">
      <cmeta:annotation_type>limitation</cmeta:annotation_type>
      <rdf:value>
        This component is only valid for temperatures above 20 degrees C
      </rdf:value>
      <dc:creator rdf:parseType="Resource">
        <vCard:N rdf:parseType="Resource">
          <vCard:Family>Doo</vCard:Family>
          <vCard:Given>Scooby</vCard:Given>
        </vCard:N>
      </dc:creator>
      <dc:date rdf:parseType="Resource">
        <dcq:dateScheme>W3C-DTF</dcq:dateScheme>
        <dcq:dateType>created</dcq:dateType>
        <rdf:value>2001-03-28</rdf:value>
      </dc:date>
    </cmeta:annotation>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 14: Recommended definition of comment and limitation annotation metadata.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:cmeta="http://www.cellml.org/metadata/1.0#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:dcq="http://purl.org/dc/qualifiers/1.0/"
  xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#">

  <rdf:Description about="#cellml_element_id">
    <cmeta:annotation rdf:parseType="Resource">
      <cmeta:annotation_type>modification</cmeta:annotation_type>
      <rdf:value>changed the equation for the sodium current</rdf:value>
      <dc:creator rdf:parseType="Resource">
        <vCard:N rdf:parseType="Resource">
          <vCard:Family>PowerPuff</vCard:Family>
          <vCard:Given>Bubbles</vCard:Given>
        </vCard:N>
      </dc:creator>
      <dc:date rdf:parseType="Resource">
        <dcq:dateScheme>W3C-DTF</dcq:dateScheme>
        <dcq:dateType>created</dcq:dateType>
        <rdf:value>2001-04-01</rdf:value>
      </dc:date>
    </cmeta:annotation>
    <cmeta:annotation rdf:parseType="Resource">
      <cmeta:annotation_type>validation</cmeta:annotation_type>
      <rdf:value>Physiome level 2</rdf:value>
      <dc:creator rdf:parseType="Resource">
        <vCard:N rdf:parseType="Resource">
          <vCard:Family>Too</vCard:Family>
          <vCard:Given>Shaggy</vCard:Given>
        </vCard:N>
      </dc:creator>
      <dc:date rdf:parseType="Resource">
        <dcq:dateScheme>W3C-DTF</dcq:dateScheme>
        <dcq:dateType>created</dcq:dateType>
        <rdf:value>2001-03-28</rdf:value>
      </dc:date>
    </cmeta:annotation>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 15: Recommended definition of modification and validation annotation metadata.

2.8 Biological Entity

This area of the metadata will almost certainly be expanded in future versions of CellML. For now, it is simply a name or database unique identifier for a biological entity, such as an ion channel, signaling pathway, or specific cell type, that is represented by the model or model component. A given CellML element can represent multiple biological entities, either as a complete group or as a list of alternatives. A CellML element that represents a list of alternative biological entities would probably be a “superclass” component, that will re-used multiple times in a model, each time to represent a different entity on the list of alternatives. For instance, a modeller might define a general “calcium-binding protein” component, and then re-use this component three times in his/her model: once to represent calmodulin, once to represent troponin C, and once to represent parvalbumin. [Note that the component re-use capabilities are not yet defined in CellML. They will be a part of the a future version of CellML. The list of alternative biological entities metadata construct is provided now for use in the future.]

Biological entity metadata is defined using a CellML-specific element, **<cmeta:bio_entity>**. A biological entity may be identified by name, database identifier, or both. Multiple database identifiers may be provided, but all except one must be marked “alternative”. The name of the biological entity is defined exactly as alternative names for CellML elements are defined (with the **<dc:title>** element, which may be qualified by a **<dcq:titleType>** element).

Each database identifier is stored in a **<cmeta:identifier>** element, which must be qualified by a **<cmeta:identifier_scheme>** element that identifies the database. The CellML metadata specification will control names for certain encoding schemes (see below). The **<cmeta:identifier>** element may also be qualified by a **<cmeta:identifier_type>** element. This element should have a value of “alternative” for all **<cmeta:identifier>** elements except for one, which is considered the primary identifier. This addresses a concern about allowing multiple database identifiers that might actually refer to different biological entities. Such an error may still occur, but marking all identifiers except one as “alternative” provides software a method by which to determine which identifier should be given precedence.

The CellML metadata specification will define the following encoding schemes:

- SWISS-PROT (SWISS-PROT protein database)
- GenBank (GenBank nucleic acid database)
- URI (URI for a web resource providing info about the biological entity)

Model authors and authors of processing software are free to define additional encoding schemes, by putting the **<entity_scheme>** element in an application-specific namespace. However, software claiming to be “CellML metadata compliant” is not required to recognize these schemes.

RDF containers can be used to indicate that a given CellML element is relevant for more than one biological entity. An **<rdf:Bag>** element can be used to indicate that the CellML element is relevant for an entire group of biological entities. An **<rdf:Alt>** element can be used to indicate that the CellML element can be relevant for one member of a group of entities. Note that the first member listed in the **<rdf:Alt>** element will be considered the preferred value. The use of the **<rdf:Bag>** element is shown in Figure 16. The use of the **<rdf:Alt>** element would be identical. “CellML metadata compliant” software will be required to recognize RDF containers in biological entity metadata. The use of RDF containers is preferred to simply repeating the **<cmeta:bio_entity>** element because it removes all ambiguity about how the group of biological entities relates to the referenced CellML element.

Figure 16 demonstrates the definition of biological entity metadata.

2.9 Copyright

The copyright metadata refers to the copyright that protects the CellML document, model, model component, or other CellML element. It is defined using the Dublin Core rights element (**<dc:rights>**), and

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:cmeta="http://www.cellml.org/metadata/1.0#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:dcq="http://purl.org/dc/qualifiers/1.0/">

  <rdf:Description about="#cellml_element_id">
    <cmeta:bio_entity>
      <rdf:Bag>
        <rdf:li rdf:parseType="Resource">
          <dc:title>calmodulin</dc:title>
          <dc:title rdf:parseType="Resource">
            <dcq:titleType>alternative</dcq:titleType>
            <rdf:value>CaM</rdf:value>
          </dc:title>
          <cmeta:identifier rdf:parseType="Resource">
            <cmeta:identifier_scheme>SWISS-PROT</cmeta:identifier_scheme>
            <rdf:value>CALM_HUMAN</rdf:value>
          </cmeta:identifier>
        </rdf:li>
        <rdf:li rdf:parseType="Resource">
          <dc:title>troponin C</dc:title>
        </rdf:li>
        <rdf:li rdf:parseType="Resource">
          <cmeta:identifier rdf:parseType="Resource">
            <cmeta:identifier_scheme>SWISS-PROT</cmeta:identifier_scheme>
            <rdf:value>PRVA_HUMAN</rdf:value>
          </cmeta:identifier>
        </rdf:li>
      </rdf:Bag>
    </cmeta:bio_entity>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 16: Recommended definition of biological entity metadata. The referenced CellML element represents the following group of proteins: calmodulin, troponin C, and parvalbumin (the protein identified by SWISS-PROT entry PRVA_HUMAN). The calmodulin biological entity has an alternative name and a database entry. The troponin C biological entity is only identified by name. The PRVA_HUMAN protein is only identified by database reference.

therefore, a given CellML element can technically have multiple copyrights. However, the recommended practice is to include only one copyright for any given element.

Figure 17 demonstrates the definition of the copyright metadata.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/">

  <rdf:Description about="#cellml_element_id">
    <dc:rights>Physiome Sciences, 2000</dc:rights>
  </rdf:Description>
</rdf:RDF>
```

FIGURE 17: Recommended definition of copyright metadata

2.10 Publisher

The publisher is the person or organization responsible for providing the model, model component, or other CellML element. A given CellML element can have multiple publishers.

Publisher metadata is defined with the Dublin Core publisher element (**<dc:publisher>**), as shown in Figure 18. Multivalued publisher metadata is handled exactly as multivalued model builder metadata. Simple repetition of the element indicates that the people or organizations publish the resource independently. The use of RDF bag (**<rdf:Bag>**) or sequence (**<rdf:Seq>**) containers indicates that the people or organization publish the resource as an unordered or ordered group, respectively.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/">

  <rdf:Description about="">
    <dc:publisher>
      University of Auckland, Bioengineering Research Group
    </dc:publisher>
  </rdf:Description>
</rdf:RDF>
```

FIGURE 18: Recommended definition of publisher metadata. Note that the empty **about** attribute indicates that this metadata refers to the CellML *document* (as opposed to the model or a specific element in the model).

2.11 Mathematical Problem Type

The mathematical problem type is a classification of the type of problem encoded in the math associated with the model or model component. It should be specified using some sort of controlled vocabulary, such

as the NIST's [GAMS classification tree](http://gams.nist.gov/)³.

Mathematical problem type is defined using a CellML-specific element, **<cmeta:math_problem>**. (Note that earlier documents use an element called **<math_problem_type>**. The name of the element has been changed to avoid confusion with the **type** element qualifier.) This element is qualified by an encoding scheme element (**<cmeta:math_problem_scheme>**), which provides the name of the controlled vocabulary used for the classification. The allowed values of this element are themselves a vocabulary controlled by the CellML RDF schema. Currently, the only allowed value is “GAMS”, which indicates that the math problem classification is taken from the NIST's [GAMS classification tree](http://gams.nist.gov/)⁴. Modellers are free to use a different controlled vocabulary for the math problem classifications by placing the **<math_problem_scheme>** element in an application-specific namespace. However, CellML metadata compliant software is not required to recognize any classification scheme other than the GAMS tree.

Figure 19 shows the recommended definition of mathematical problem type metadata.

```
<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:cmeta="http://www.cellml.org/metadata/1.0#"

  <rdf:Description about="#cellml_element_id">
    <cmeta:math_problem rdf:parseType="Resource">
      <cmeta:math_problem_scheme>GAMS</cmeta:math_problem_scheme>
      <rdf:value>I1a</rdf:value>
    </cmeta:math_problem>
  </rdf:Description>
</rdf:RDF>
```

FIGURE 19: Recommended definition of the mathematical problem type metadata. The meaning of the value “GAMS” for the encoding scheme will be controlled by the CellML metadata specification.

2.12 Contributor

Contributor metadata indicates that a person contributed to a resource, but did not actually create it (an example of this is an editor).

Contributor metadata is defined using the Dublin Core contributor element, **<dc:contributor>**, as shown in Figure 20. Multivalued contributor metadata is handled exactly as multivalued model builder metadata. Simple repetition of the element indicates that the people contributed to the resource independently. The use of RDF bag (**<rdf:Bag>**) or sequence (**<rdf:Seq>**) containers indicates that the people contributed to the resource as an unordered or ordered group, respectively.

2.13 Description

Description metadata is a short description of the referenced resource.

Description metadata is defined with the Dublin Core description element, **<dc:description>**. This element is qualified by a type element (**<dcq:descriptionType>**), which may have a value of “abstract” or “table of contents”. The “abstract” type will probably be most common in CellML metadata.

Figure 21 shows how to define description metadata.

³<http://gams.nist.gov/>

⁴<http://gams.nist.gov/>

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#">

  <rdf:Description about="#cellml_element_id">
    <dc:contributor rdf:parseType="Resource">
      <vCard:N rdf:parseType="Resource">
        <vCard:Family>Flinstone</vCard:Family>
        <vCard:Given>Fred</vCard:Given>
      </vCard:N>
    </dc:contributor>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 20: Recommended definition of contributor metadata.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:dcq="http://purl.org/dc/qualifiers/1.0/">

  <rdf:Description about="#cellml_element_id">
    <dc:description rdf:parseType="Resource">
      <dcq:descriptionType>abstract</dcq:descriptionType>
      <rdf:value>
        This element uses simple mass-action kinetics to describe the
        A + B <-> C + D reaction.
      </rdf:value>
    </dc:description>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 21: Recommended definition of the Dublin Core description metadata.

2.14 Metadata about People

Warren and Melanie were confident that *someone* must have solved this problem before. After all, the need to store information about people is common to most projects. Perhaps it is so common that everyone just defines their own method, because the only existing RDF definition of metadata about people is a note submitted to the W3C in February 2001 entitled [Representing vCard Objects in RDF/XML](http://www.w3.org/TR/2001/NOTE-vcard-rdf-20010222)⁵. (This note is the work of Renato Iannella working at the Distributed Systems Technology Centre at the University of Queensland and originally appeared on their [RDF project page](http://www.dstc.edu.au/Research/Projects/rdf/)⁶.) This leaves us with two choices: use the “vCard in RDF” option, or define our own metadata elements for people.

The vCard data model includes all of the information that the [original CellML requirements](http://www.cellml.org/private/fundamentals/requirements.html)⁷ indicated we need. This is:

- Name, split into last name, first name, and middle name/initial
- Contact info, which may include mailing address(es), e-mail address(es), and phone number(s)
- Affiliation

It is an attractive option to use an existing data model when (1) the model is complete, meaning that it includes all of the information we need or (2) we will use all or almost all of the elements in the model, even if we need to add some additional elements. Furthermore, the existence of an RDF implementation of a data model makes it more attractive for use in CellML. The vCard data model meets the first option, and also has an existing RDF implementation. Therefore, we should use it.

However, the vCard data model includes some elements that are not necessary for CellML metadata, such as nickname and birthday. We will therefore not require CellML processing software to recognize those elements. However, model authors are free to use them. That is, the use of vCard elements outside of the list defined in the CellML metadata specification will not invalidate the metadata, but these elements may not necessarily be recognized by all CellML metadata compliant processing software.

The CellML metadata specification should recommend the use of the following “vCard in RDF” elements to meet the information needs in the requirements:

- **<vCard:N>** (the name construct), with all of its subelements:
 - **<vCard:Family>**: the person’s family, or last name
 - **<vCard:Given>**: the person’s given, or first name
 - **<vCard:Other>**: additional names, used for middle names and initials
 - **<vCard:Prefix>**: honorific prefixes, such as “Dr.”
 - **<vCard:Suffix>**: suffixes such as “III” and “Jr.”
- **<vCard:ADR>** (the mailing address construct), with all of its subelements:
 - **<vCard:Pobox>**: post office box
 - **<vCard:Street>**: street address
 - **<vCard:Locality>**: city, town, rural route, etc.
 - **<vCard:Region>**: state, etc.
 - **<vCard:Country>**: country
 - **<vCard:Pcode>**: postal code (such as the American zip code)

⁵<http://www.w3.org/TR/2001/NOTE-vcard-rdf-20010222>

⁶<http://www.dstc.edu.au/Research/Projects/rdf/>

⁷<http://www.cellml.org/private/fundamentals/requirements.html>

- **<vCard:Extadd>**: extended address field. This is used to include the company or institution name.
- **<vCard:EMAIL>** (the e-mail address construct)
- **<vCard:TEL>** (the telephone number construct)
- **<vCard:ORG>** (the organization construct, which maps to the CellML requirement to be able to store a person’s affiliation), with all of its subelements:
 - **<vCard:Orgname>**: the name of the organization (i.e., “The University of Auckland”)
 - **<vCard:Orgunit>**: the division or department (i.e., “The Bioengineering Research Group”)
- **<vCard:TITLE>**: the person’s job title (not required in original requirements list, but deemed useful enough to include in the metadata specification).
- **<vCard:ROLE>**: the person’s job role (not required in original requirements list, but deemed useful enough to include in the metadata specification).

The **<rdf:type>** element is used to specify “type parameters” on certain vCard elements. For instance, an address may be typed as domestic, international, postal, parcel, home, work, or preferred. Note that one address may be given more than one type. See section 3.3 of the [vCard in RDF document](#)⁸ for more info.

In addition, lists of alternative values for some vCard elements are implemented using the **<rdf:Alt>** container. Ordered groups of values are implemented using the **<rdf:Seq>** container, and unordered groups of values are implemented using the **<rdf:Bag>** container. See section 3.2 of the [vCard in RDF document](#)⁹ for more info. The CellML metadata specification will only require processing software to recognize the following group values:

- **<rdf:Seq>** and **<rdf:Bag>** containers for job title and job role metadata.
- **<rdf:Alt>** and **<rdf:Bag>** containers for organization metadata.
- **<rdf:Seq>** for organization units, to support the implied ordering (see special case 2 in section 3.4 of the [vCard in RDF document](#).)

The metadata specification will specifically NOT require the use of the **<rdf:Alt>** container for multivalued contact info. Modellers will instead be recommended to use the **<rdf:type>** element with repeated instances of the basic vCard contact info elements. (See section 3.3 of the [vCard in RDF document](#).)

2.15 Examples of the Use of vCard in CellML Metadata

Figure 22 shows the use of vCard to supply basic information about a model builder.

Figure 23 shows the use of vCard to supply more detailed information about a model builder. This example attempts to show many of the complicated uses of vCard. It is unlikely that real world examples will be this pathological.

The example demonstrates how to include multiple e-mail addresses, phone numbers, and mailing addresses for a person, and how to indicate the type of each piece of contact info. For instance, Fred Flinstone’s preferred e-mail address is his work e-mail address, `fred@bedrock.edu`. The example also demonstrates the use of an **<rdf:Seq>** container to provide multiple subunits in the organization unit metadata. The metadata in this example is correctly interpreted to say that Dr. Fred Flinstone works in

⁸<http://www.w3.org/TR/2001/NOTE-vcard-rdf-20010222>

⁹<http://www.w3.org/TR/2001/NOTE-vcard-rdf-20010222>

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#">

  <rdf:Description about="#cellml_element_id">
    <dc:creator rdf:parseType="Resource">
      <vCard:N rdf:parseType="Resource">
        <vCard:Family>Powerpuff</vCard:Family>
        <vCard:Given>Bubbles</vCard:Given>
      </vCard:N>
      <vCard:EMAIL>
        <rdf:value>bubbles@townville.net</rdf:value>
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#internet" />
      </vCard:EMAIL>
      <vCard:TITLE>PowerPuff Girl</vCard:TITLE>
      <vCard:ROLE>fighting crime before bedtime</vCard:ROLE>
    </dc:creator>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 22: Use of vCard elements to define metadata about a model builder.

the Parallel Computing Research Group, which is a subunit of the Department of Computer Science at the University of Bedford.

```

<rdf:RDF
  xmlns:rdf="http://www.w3.org/1999/02/22-rdf-syntax-ns#"
  xmlns:dc="http://purl.org/dc/elements/1.0/"
  xmlns:vCard="http://www.w3.org/2001/vcard-rdf/3.0#"
  <rdf:Description about="#cellml_element_id">
    <dc:creator rdf:parseType="Resource">
      <vCard:FN rdf:parseType="Literal">Dr. Fred Flinstone</vCard:FN>
      <vCard:N rdf:parseType="Resource">
        <vCard:Family>Flinstone</vCard:Family>
        <vCard:Given>Fred</vCard:Given>
        <vCard:Prefix>Dr.</vCard:Prefix>
      </vCard:N>
      <vCard:EMAIL rdf:parseType="Resource">
        <rdf:value>fred@bedrock.edu</rdf:value>
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#internet" />
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#pref" />
      </vCard:EMAIL>
      <vCard:EMAIL rdf:parseType="Resource">
        <rdf:value>fred.flinstone@yahoo.com</rdf:value>
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#internet" />
      </vCard:EMAIL>
      <vCard:ORG rdf:parseType="Resource">
        <vCard:Orgname>The University of Bedrock</vCard:Orgname>
        <vCard:Orgunit><rdf:Seq>
          <rdf:li>Department of Computer Science</rdf:li>
          <rdf:li>Parallel Computing Research Group</rdf:li>
        </rdf:Seq></vCard:Orgunit>
      </vCard:ORG>
      <vCard:TEL rdf:parseType="Resource">
        <rdf:value>609-999-1111</rdf:value>
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#work" />
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#voice" />
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#pref" />
      </vCard:TEL>
      <vCard:TEL rdf:parseType="Resource">
        <rdf:value>609-777-7777</rdf:value>
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#cell" />
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#voice" />
      </vCard:TEL>
      <vCard:ADR rdf:parseType="Resource">
        <rdf:value rdf:parseType="Resource">
          <vCard:Street>800 Paleolithic Drive</vCard:Street>
          <vCard:Locality>Bedrock</vCard:Locality>
          <vCard:Pcode>6767</vCard:Pcode>
          <vCard:Country>Australia</vCard:Country>
        </rdf:value>
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#work" />
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#pref" />
      </vCard:ADR>
      <vCard:ADR rdf:parseType="Resource">
        <rdf:value rdf:parseType="Resource">
          <vCard:Street>16 Yabba Dabba Doo Ave.</vCard:Street>
          <vCard:Locality>Bedrock</vCard:Locality>
          <vCard:Pcode>6767</vCard:Pcode>
          <vCard:Country>Australia</vCard:Country>
        </rdf:value>
        <rdf:type rdf:resource="http://imc.org/vCard/3.0#home" />
      </vCard:ADR>
    </dc:creator>
  </rdf:Description>
</rdf:RDF>

```

FIGURE 23: Use of vCard elements to define detailed metadata about a model builder. See text for more info.
