EMBL-EBI

European Bioinformatics Institute,
an Outstation of the
European Molecular Biology Laboratory

MIRIAM

MIRIAM Resources:
a robust annotation and
cross referencing framework

Nick Juty, Nicolas Le Novère, Camille Laibe

Introduction

More than ever, the era of *data integration* has highlighted the key requirement to reference specific data in an unambiguous and perennial way, in order to enable community-level sharing, development, exchange and reuse of information. In the field of Systems Biology, which is concerned with creating quantitative models of biological processes, these requirements have directly led to the creation of the Minimal Information Required In the Annotation of Models (MIRIAM, http://biomodels.net/miriam/) [1]. MIRIAM provides a specific set of guidelines that can be implemented within any structured modelling format.
To support the implementation of MIRIAM guidelines, we have developed MIRIAM Resources (http://www.ebi.ac.uk/miriam/) [2]. This software infrastructure facilitates the use of controlled annotations in the identification of model components. The annotation procedures described in the MIRIAM standard entail the use of a *Uniform Resource Identifier* (URI), thereby uniquely identifying a component without reference to its physical location.

MIRIAM Standard

Qualifiers

Annotations

Annotations are essential for data identification and semantics:

- data understanding and reuse
- data comparison
- data integration

Cross-references

Characteristics of a useful identifier for cross-references:

- unique and not ambiguous
- perennial
- standard compliant
- resolvable
- free of use

SBML sample

```
[...]
<species metaid="metaid_0000006"
         id="L_EGFR"
         compartment="compartment"
         initialConcentration="0">
  <annotation>
    <rdf:RDF xmlns:rdf="http://www.w3.org/1999/02/22rdfsyntaxns#"
             xmlns:bqbiol="http://biomodels.net/biologyqualifiers/">
      <rdf:Description rdf:about="#metaid_0000006">
        <bqbiol:hasPart>
          <rdf:Bag>
            <rdf:li rdf:resource="urn:miriam:uniprot:P07522" />
            <rdf:li rdf:resource="urn:miriam:uniprot:Q9QX70_RAT" />
          </rdf:Bag>
        </bqbiol:hasPart>
      </rdf:Description>
    </rdf:RDF>
  </annotation>
</species>
[...]
```
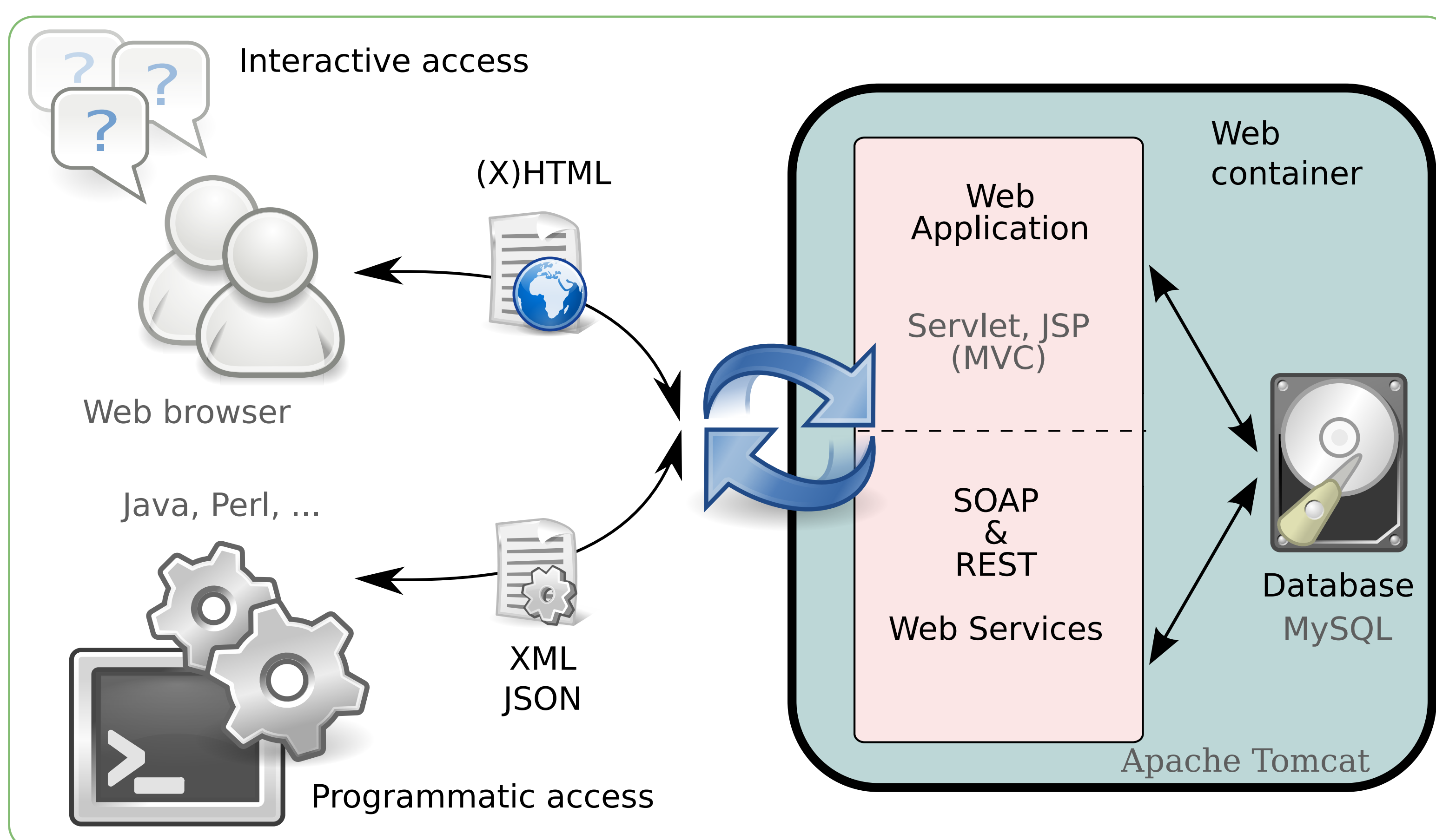
MIRIAM URI

Human calmodulin: P62158 in UniProt - urn:miriam:uniprot:P62158
Alcohol dehydrogenase: 1.1.1.1 in EC code - urn:miriam:ec-code:1.1.1.1
Activation of MAPKK activity: GO:0000186 in Gene Ontology - urn:miriam:obo.go:GO%3A0000186

Architecture

Resources

MIRIAM Resources are a set of on-line services centered around a catalogue of data types. Data types can be ontologies, such as Gene Ontology, or primary data resources available via the Web, such as UniProt and PubMed. Each of these is uniquely identified within the MIRIAM database, and information stored regarding the corresponding physical URLs (data resources) through which their associated data can be accessed. Conversion services, to generate a *MIRIAM URI* from a data type name and dataset identifier, and resolution services, to generate physical locations for a given MIRIAM URI, are provided through SOAP and REST Web Services.
MIRIAM annotations have become widely accepted and now used and supported by various communities, such as BioModels.net [3], SBML, CellML and BioPAX.The utility of MIRIAM URIs is not restricted solely to the field of quantitative modelling; MIRIAM Resources are now used by the Proteomics Standards Initiative (PSI) [4].
MIRIAM Resources allow easy access to MIRIAM URIs, as well as associated information, and facilitate the identification, exchange, and integration of data in the Life Sciences. The source code for the entire project is available, under the GNU General Public License, at: http://sourceforge.net/projects/miriam/.
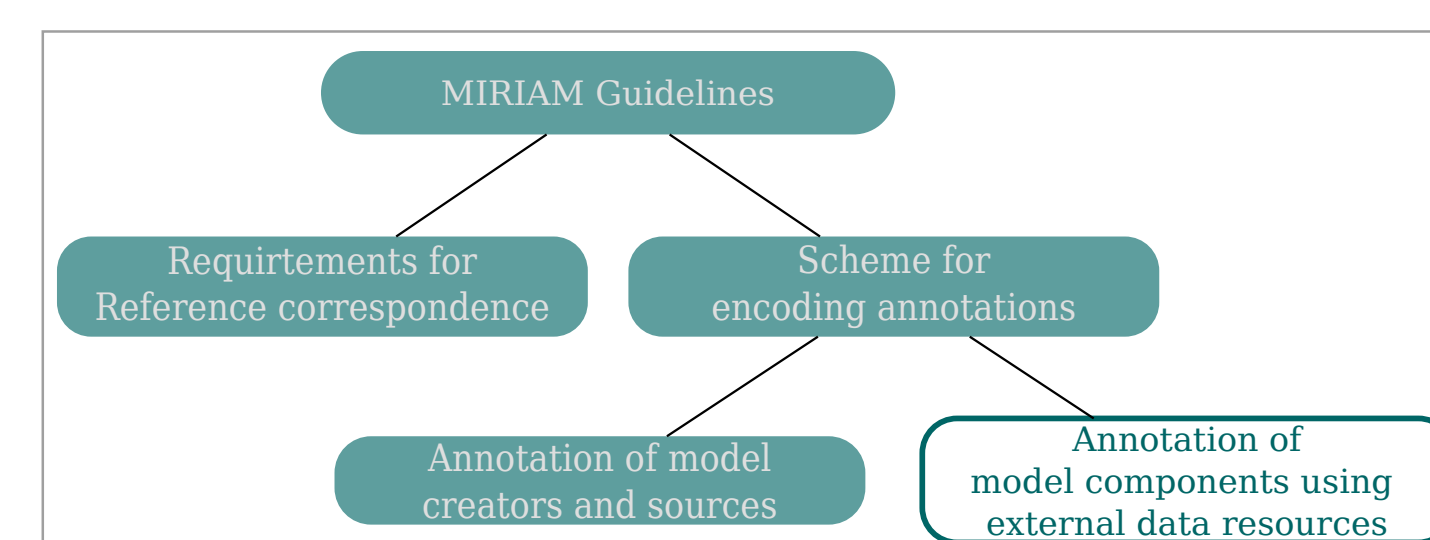

Features

- Catalogue of data types (open submission)
  - General information
  - Resources
  - Tags
  - Examples of usage (SBML, CellML, BioPAX, ...)
  - Web Services provided
- Web Services (SOAP and REST)
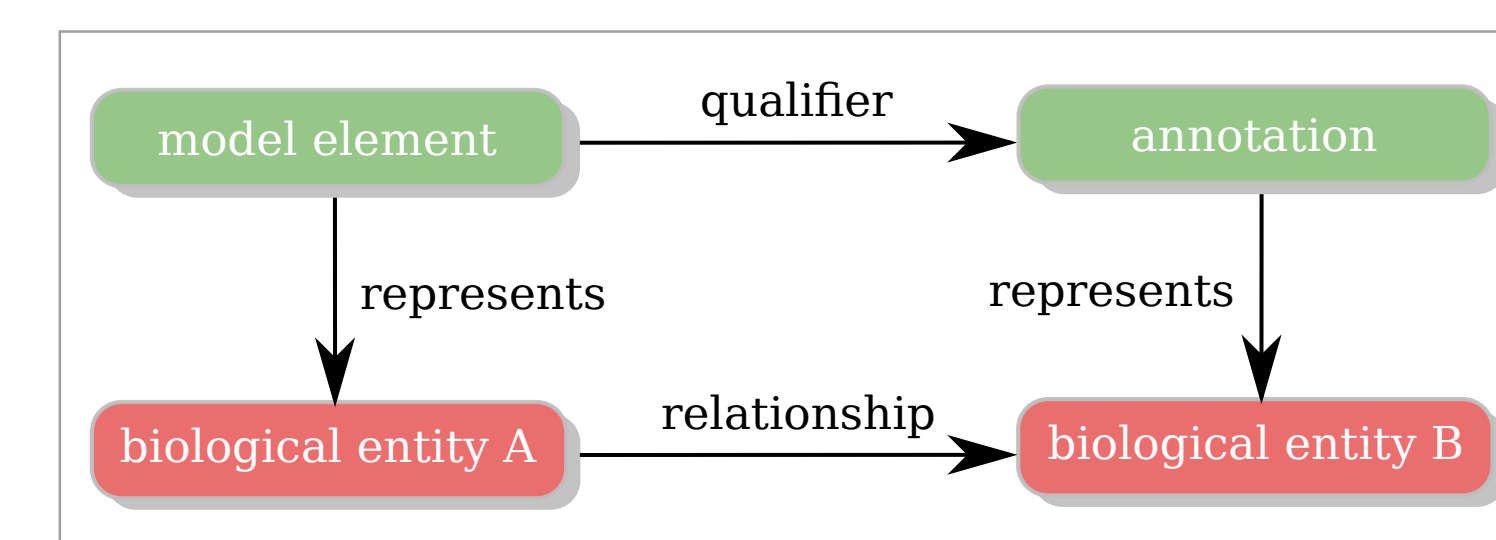- Exports (XML)
- Daily resources health check

References

[1] N. Le Novère, A. Finney, M. Hucka, U. Bhalla, F. Campagne, J. Collado-Vides, E.J. Crampin, M. Halstead, E. Klipp, P. Mendes, P. Nielsen, H. Sauro, B. Shapiro, J.L. Snoep, H.D. Spence, and B.L. Wanner. **Minimum Information Requested In the Annotation of biochemical Models (MIRIAM)**. *Nature Biotechnology*, 23(12):1509–1515, 2005.
[2] C. Laibe and N. Le Novère. **MIRIAM Resources: tools to generate and resolve robust cross-references in Systems Biology**. *BMC Systems Biology*, 1: 58, 2007.
[3] Le Novère N. **BioModels.net, tools and resources to support Computational Systems Biology**. Proceedings of the 4th Workshop on Computation of Biochemical Pathways and Genetic Networks. Logos, Berlin, pp. 69-74, 2005.
[4] Orchard, S., Jones, P., Taylor, C., Zhu, W., Julian, R.K., Jr., Hermjakob, H. and Apweiler, R. **Proteomic data exchange and storage: the need for common standards and public repositories**. Methods Mol Biol 367: 261-270, 2007.

laibe@ebi.ac.uk, juty@ebi.ac.uk
lenov@ebi.ac.uk