



Proposal for a new URI Scheme in MIRIAM

Camille Laibe

Okinawa, Jan 2008



{data type, identifier}

Two syntaxes:

- URL: `http://www.myResource.org/#myIdentifier`
- URL: `urn:lsid:myResources.org:myIdentifier`



- <http://biomodels.net/MIRIAM/>
- <http://www.ebi.ac.uk/IntEnz/> (obsolete)
- <http://www.ec-code.org/>
- <urn:lsid:ec-code.org>
- <http://www.ebi.ac.uk/chebi/>
- <urn:lsid:uniprot.org:uniprot>
- <http://www.taxonomy.org/>





Confusion:

- identifier of an abstract concept and the physical address of a Web page

- KEGG Compound

- <http://www.genome.jp/kegg/compound/>
- [http://www.genome.jp/dbget-bin/www_bget?cpd:\\$id](http://www.genome.jp/dbget-bin/www_bget?cpd:$id)

- ChEBI

- <http://www.ebi.ac.uk/chebi/>
- [http://www.ebi.ac.uk/chebi/searchFreeText.do?searchString=\\$id](http://www.ebi.ac.uk/chebi/searchFreeText.do?searchString=$id)





Ownership:

- domain name (example: taxonomy.org)

Elements:

- domain names are not trademarks
- don't used as physical locations

But: **collisions!**



Most of the URNs used are actually LSIDs

```
urn:lsid:<authority>:<namespace>:<objectId>[:<version>]
```

“it is recommended that [the domain name in the authority section] be owned by the organization that assigns an LSID in question” (LSID specs)





- <http://www.taxonomy.org/>
- <http://www.ebi.ac.uk/chebi/>

Questions:

- Who created these URIs?
- What are they used for?
- What can one do with them?
- ...



Neutral URIs:

- <http://www.taxonomy.org/>
- <http://www.ebi.ac.uk/chebi/>

Same pattern:

- <http://www.eccode.org/>
- <http://biomodels.net/MIRIAM/>

URNs:

- <urn:lsid:uniprot.org:uniprot>
- <urn:oai:arXiv.org>



Fortunately:

There are solutions

Unfortunately:

That implies changing the “perennial” current URIs
used in MIRIAM...





- the top level part of the URI naming structure
 - http://www.en.wikipedia.org/wiki/URI_Scheme
- equivalent to the “protocol” part
- URIs look like:
 - `mir://dataType/entity`
 - `miriam:dataType:entity`





Advantages:

- consistency of the URIs
- no confusion possible with a physical address
- reference to MIRIAM included in the URIs

Disadvantages:

- need to be registered for perennality purposes
=> **lengthy** (via Internet Assigned Numbers Authority)
- URIs can't be on the Web (W3C).





- portions of URN space delegated to a specific usage
- Example:
 - urn:oasis:names:tc:ubl:schema:xsd:Invoice-1.0
- Possible URI:
 - urn:miriam:dataType:entity
- no protocol associated
- no dereference possible
- easier to register than a URI scheme





- “http:” scheme is a standard
- primarily for Web documents...
...but now it is an obsolete idea:

```
<go:term rdf:about="http://www.geneontology.org/go#GO:0000002" >
  <go:accession>GO:0000002</go:accession>
  <go:name>mitochondrial genome maintenance</go:name>
  <go:definition>The maintenance of the structure and integrity of the mitochondrial genome;
  <go:is_a rdf:resource="http://www.geneontology.org/go#GO:0007005" />
  <go:dbxref rdf:parseType="Resource">
    <go:database_symbol>InterPro</go:database_symbol>
    <go:reference>IPR009446</go:reference>
  </go:dbxref>
  <go:dbxref rdf:parseType="Resource">
    <go:database_symbol>Pfam</go:database_symbol>
    <go:reference>PF06420</go:reference>
  </go:dbxref>
</go:term>
```



Examples:

- <http://www.miriam.org/dataType/entity>
- <http://www.miriam-uri.org/id/dataType/entity>





- based on an existing, robust and widely used scheme
- consistency of the URIs
- no need for a heavy registration procedure
- possibility to be “on the Web”, if we register a domain name (optional but advised)





Contribute to the **Semantic Web**:

resolution service for the URIs

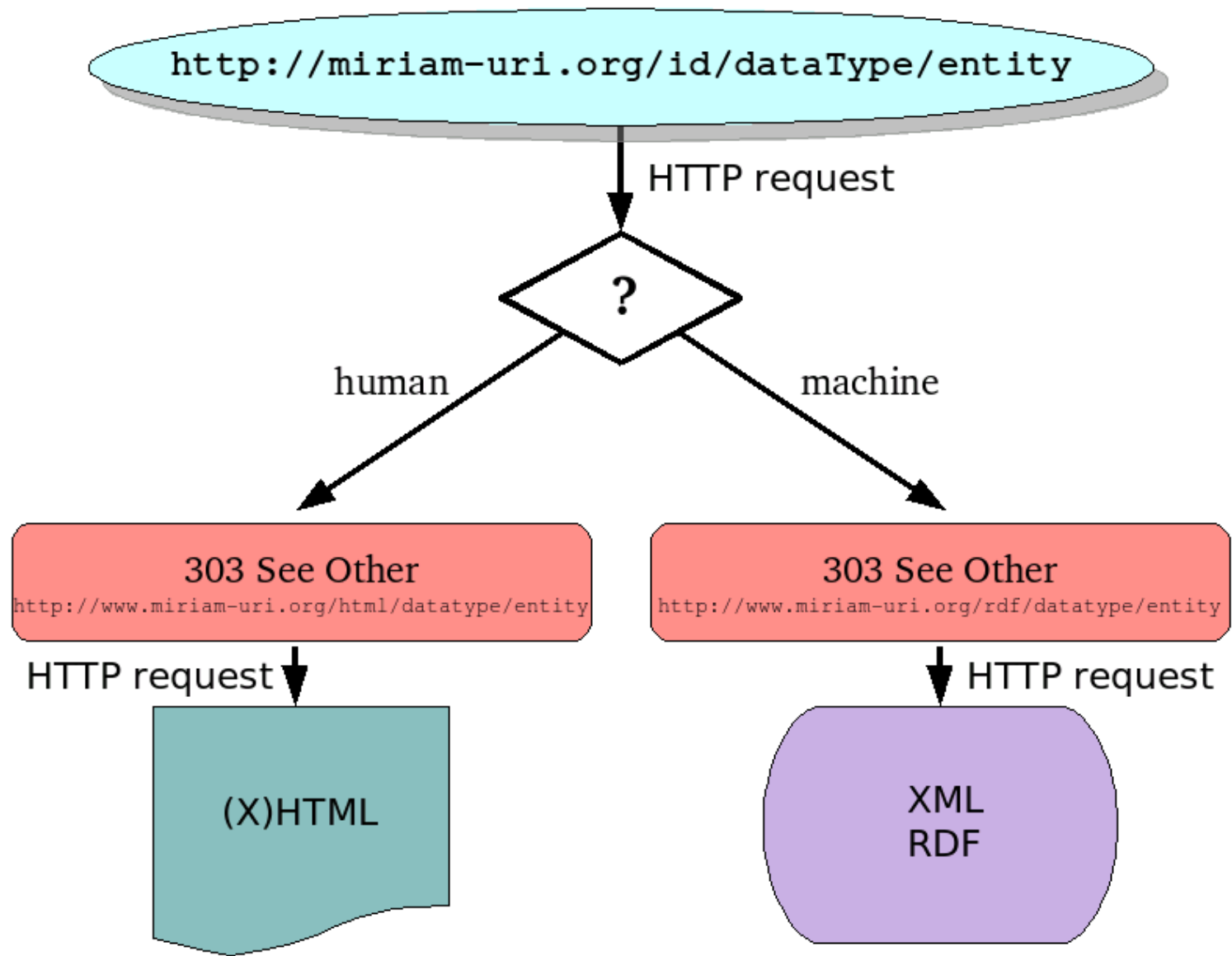
not directly to the resources

(using the 303 “See Other” response code)

agent type dependant response:

- human: (X)HTML page
- software: XML, RDF, ...





Complete document available:

http://www.ebi.ac.uk/~laibe/miriam/Miriam_URI_Scheme_Proposal.pdf



Thank you

