Identifiers.org and MIRIAM Registry: perennial identifiers for cross-referencing

Nick Juty

BioModels.net

EMBL-EBI

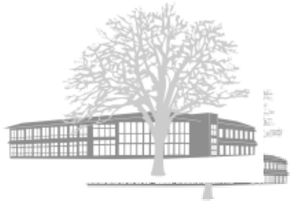
18th August, COMBINE 2012, Ontario, Canada

Outline

- Intro
  - URIs (Uniform Resource Identifiers)…

- MIRIAM Registry (namespaces)
  - usage

- Identifiers.org resolving service
  - levels of identification
  - customisation
  - usage



Outline

Intro

- URIs (Uniform Resource Identifiers)…

MIRIAM Registry (namespaces)

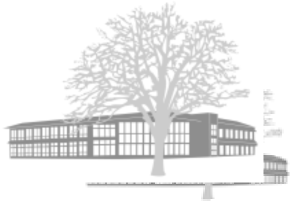
- usage

Identifiers.org resolving service

- levels of identification
- customisation
- usage

CONTRIBUTE

USE



Introduction

Computational Biology

  need to specify model components

    understanding data

    reusing data

    comparing data

    integrating data

    converting data

    linking data

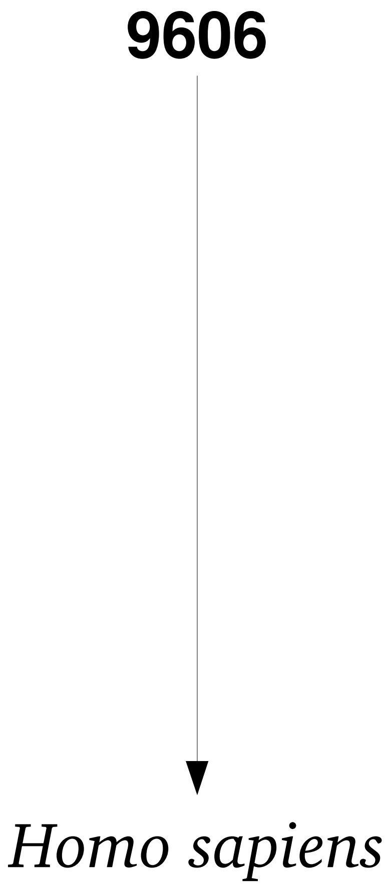
  with annotations - identifiers/xref

  - need for xrefs true of any kind of data

The need for namespaces

9606

The need for namespaces



9606

Homo sapiens

The need for namespaces

9606

Arabian Tea
(GRIN plant taxonomy)

Homo sapiens

The need for namespaces

9606

Perching bird
(BOLD taxonomy)

Arabian Tea
(GRIN plant taxonomy)

Homo sapiens
(NCBI)
Taxonomy

Xref / identifiers

Building Cross-references/identifiers

namespace – shared list (Registry)

data-set identifier – provided by collection

URI identifiers

URN location independent (not directly resolvable)

Xref / identifiers

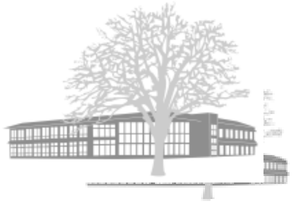
Building Cross-references/identifiers

namespace – shared list (Registry)

data-set identifier – provided by collection

URI identifiers

URN urn:miriam:ABC:12345



Xref / identifiers

Building Cross-references/identifiers

namespace – shared list (Registry)

data-set identifier – provided by collection

URI identifiers

URN urn:miriam:ABC:12345

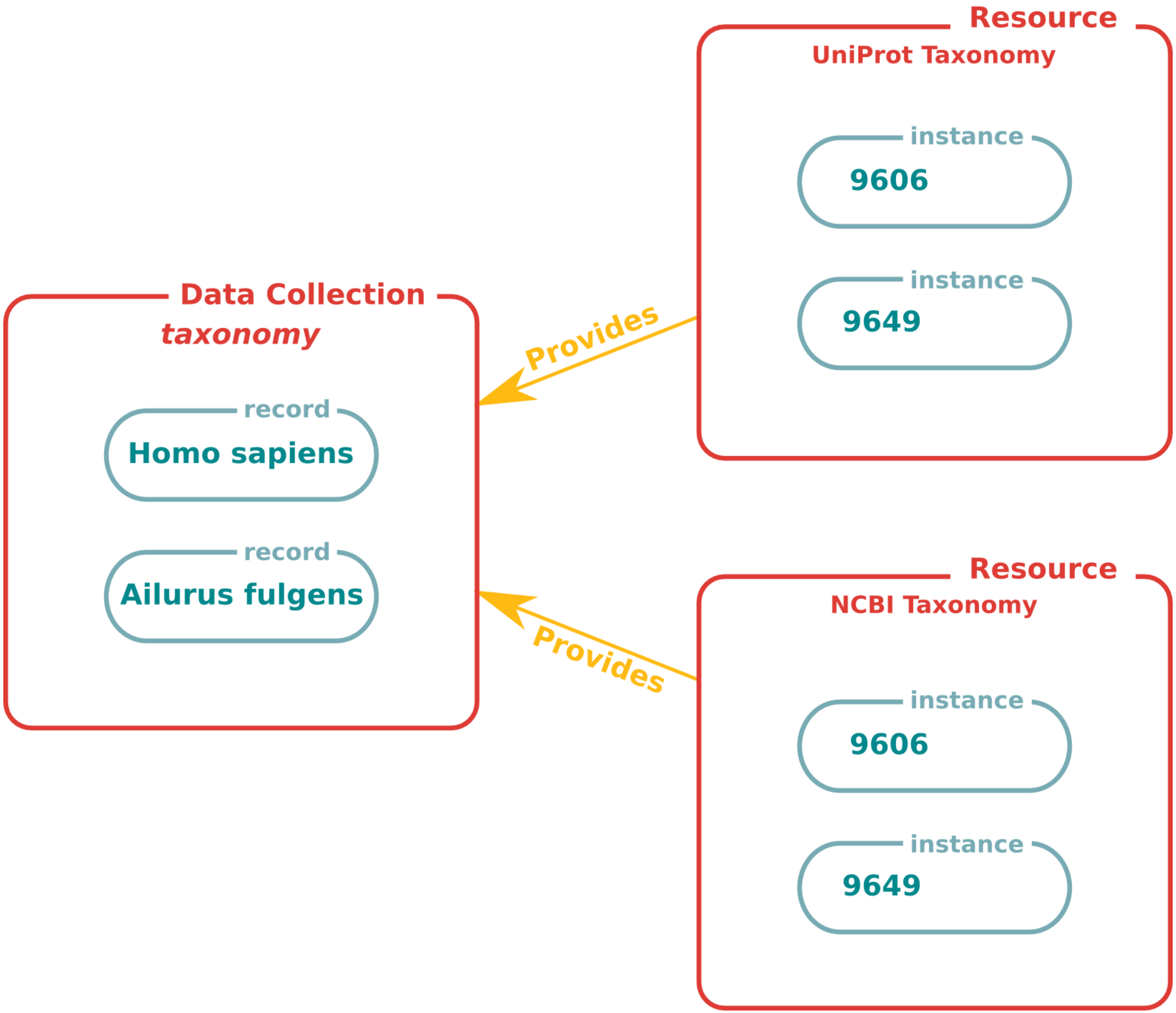
Identifiers.org: http://identifiers.org/ABC/12345

avoid URL fragility

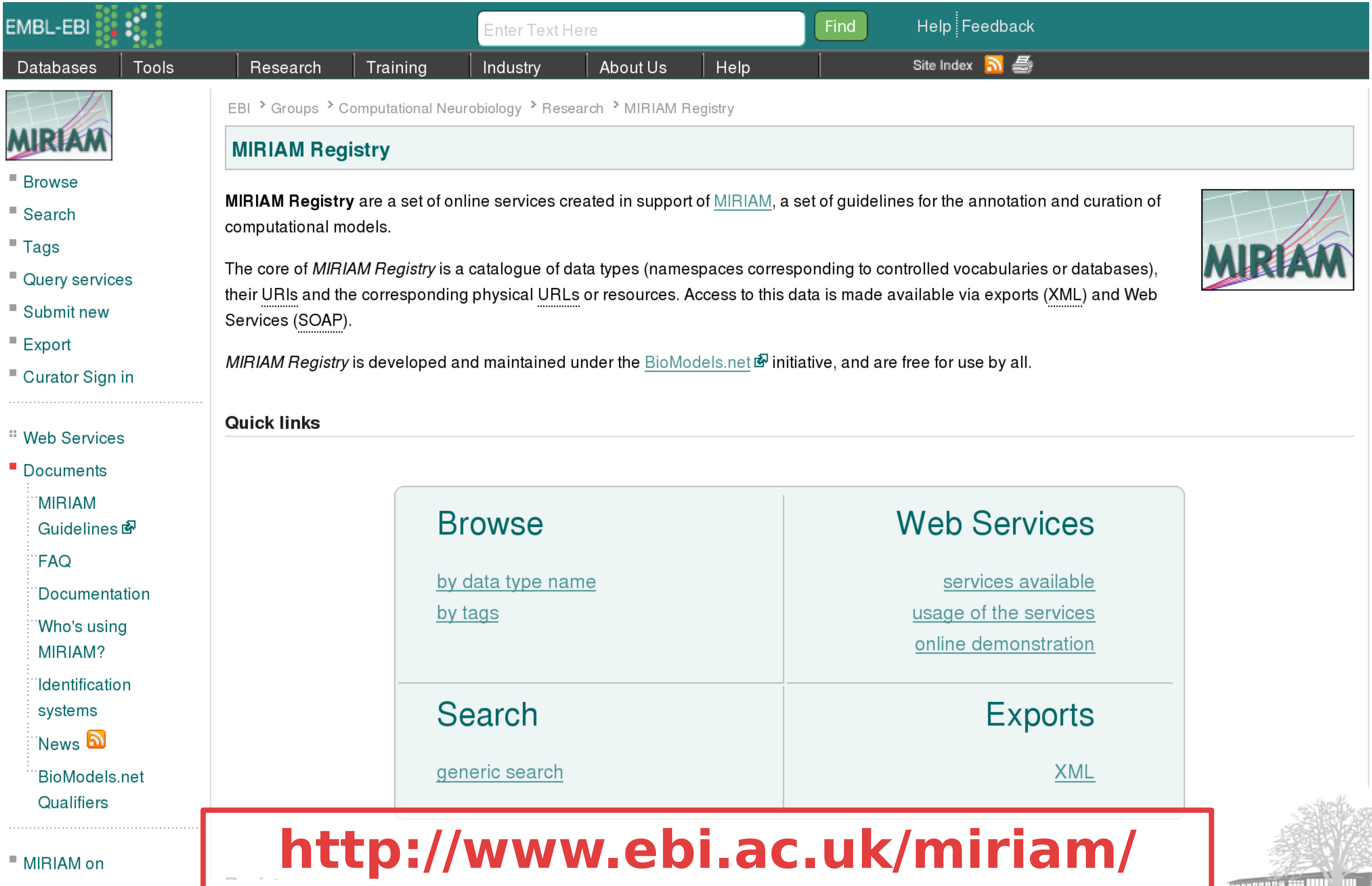
convenience of resolvability

Created as a layer above the Registry

Nomenclature



MIRIAM Registry



MIRIAM Registry

MIRIAM Registry are a set of online services created in support of MIRIAM, a set of guidelines for the annotation and curation of computational models.

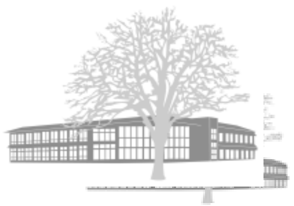
The core of *MIRIAM Registry* is a catalogue of data types (namespaces corresponding to controlled vocabularies or databases), their URIs and the corresponding physical URLs or resources. Access to this data is made available via exports (XML) and Web Services (SOAP).

MIRIAM Registry is developed and maintained under the BioModels.net initiative, and are free for use by all.

Quick links

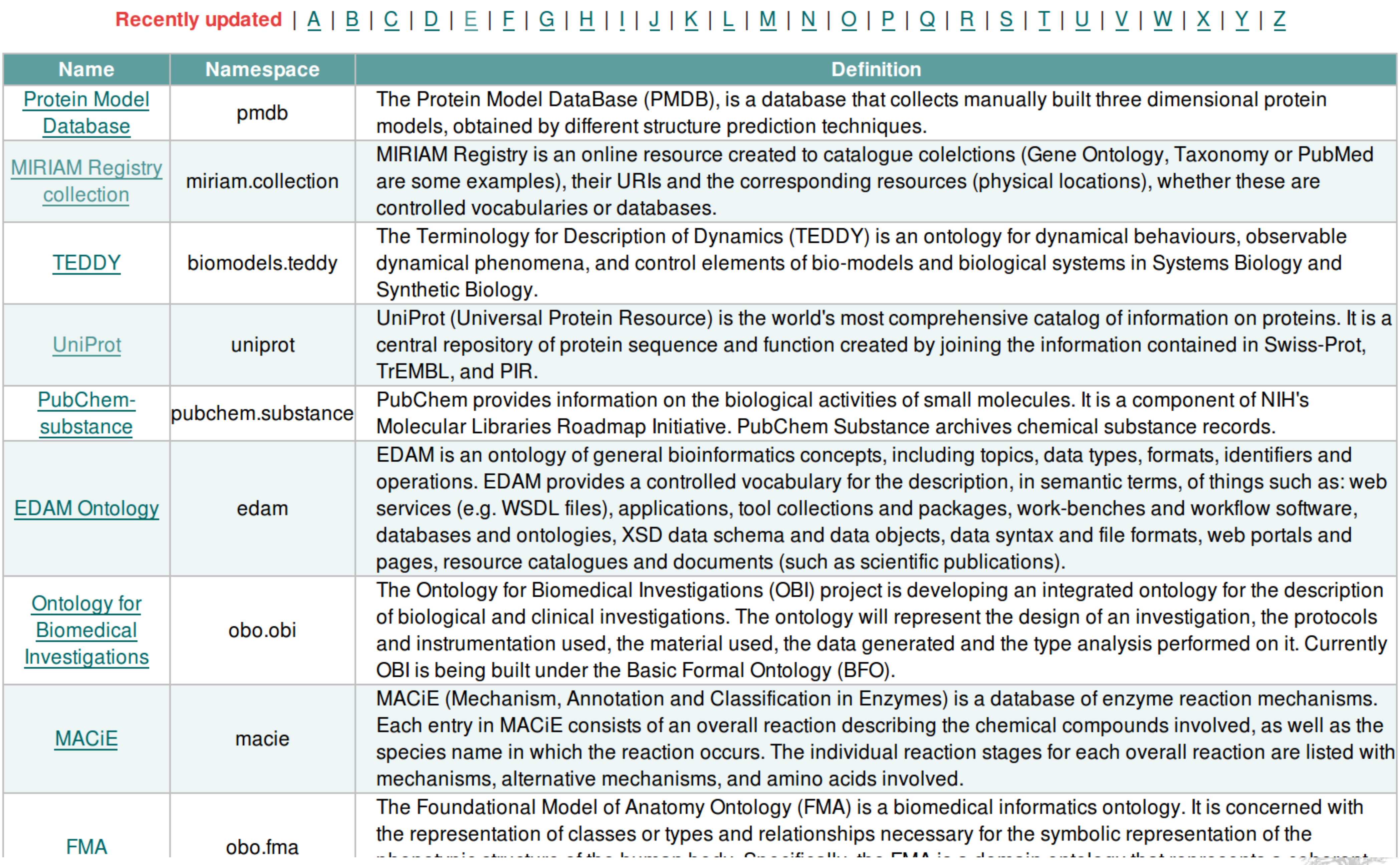
Browse by data type name by tags	**Web Services** services available usage of the services online demonstration
Search generic search	**Exports** XML

http://www.ebi.ac.uk/miriam/



EMBL-EBI

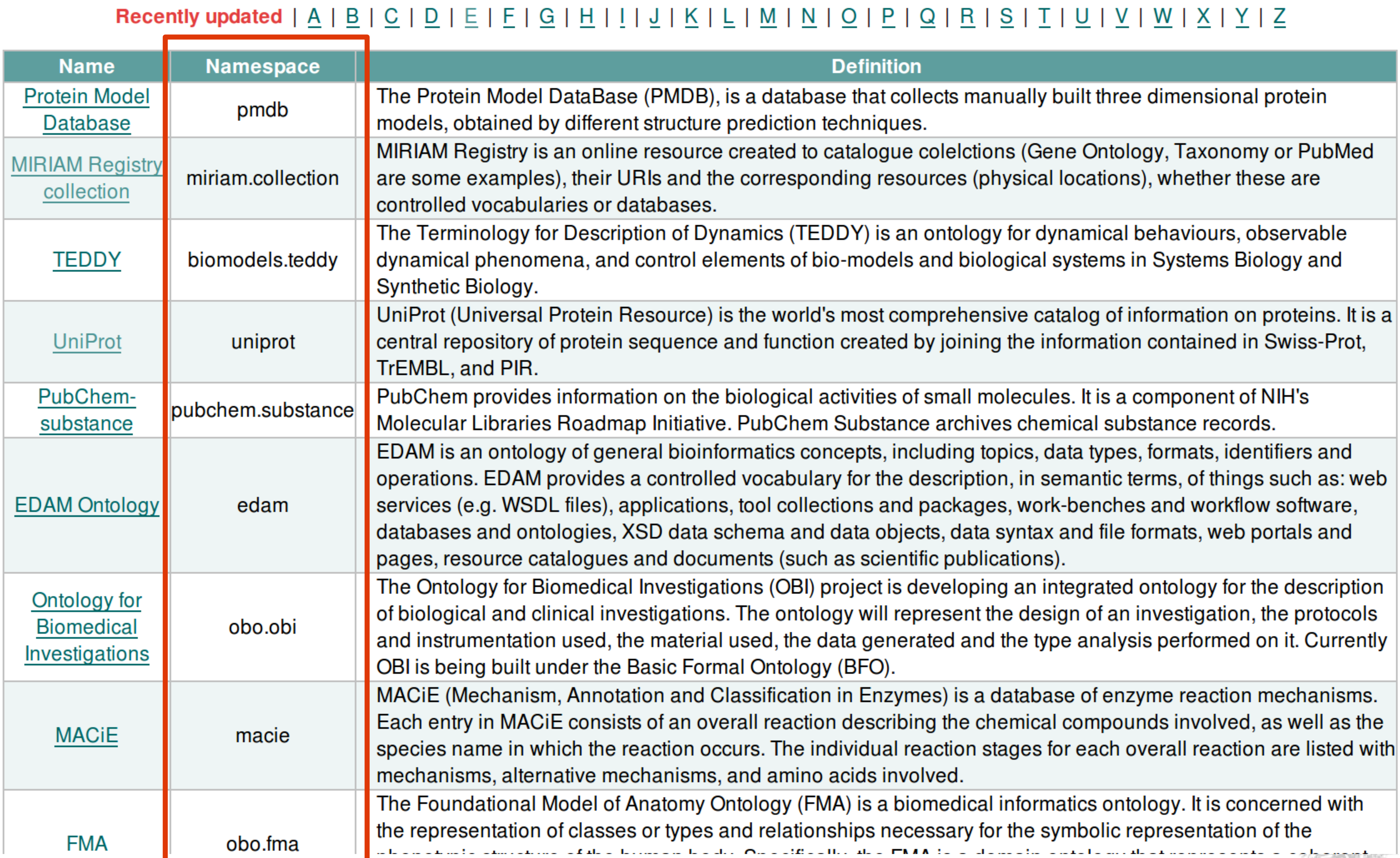
List of data collections



Recently updated | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z

Name	Namespace	Definition
Protein Model Database	pmdb	The Protein Model DataBase (PMDB), is a database that collects manually built three dimensional protein models, obtained by different structure prediction techniques.
MIRIAM Registry collection	miriam.collection	MIRIAM Registry is an online resource created to catalogue colelctions (Gene Ontology, Taxonomy or PubMed are some examples), their URIs and the corresponding resources (physical locations), whether these are controlled vocabularies or databases.
TEDDY	biomodels.teddy	The Terminology for Description of Dynamics (TEDDY) is an ontology for dynamical behaviours, observable dynamical phenomena, and control elements of bio-models and biological systems in Systems Biology and Synthetic Biology.
UniProt	uniprot	UniProt (Universal Protein Resource) is the world's most comprehensive catalog of information on proteins. It is a central repository of protein sequence and function created by joining the information contained in Swiss-Prot, TrEMBL, and PIR.
PubChem-substance	pubchem.substance	PubChem provides information on the biological activities of small molecules. It is a component of NIH's Molecular Libraries Roadmap Initiative. PubChem Substance archives chemical substance records.
EDAM Ontology	edam	EDAM is an ontology of general bioinformatics concepts, including topics, data types, formats, identifiers and operations. EDAM provides a controlled vocabulary for the description, in semantic terms, of things such as: web services (e.g. WSDL files), applications, tool collections and packages, work-benches and workflow software, databases and ontologies, XSD data schema and data objects, data syntax and file formats, web portals and pages, resource catalogues and documents (such as scientific publications).
Ontology for Biomedical Investigations	obo.obi	The Ontology for Biomedical Investigations (OBI) project is developing an integrated ontology for the description of biological and clinical investigations. The ontology will represent the design of an investigation, the protocols and instrumentation used, the material used, the data generated and the type analysis performed on it. Currently OBI is being built under the Basic Formal Ontology (BFO).
MACiE	macie	MACiE (Mechanism, Annotation and Classification in Enzymes) is a database of enzyme reaction mechanisms. Each entry in MACiE consists of an overall reaction describing the chemical compounds involved, as well as the species name in which the reaction occurs. The individual reaction stages for each overall reaction are listed with mechanisms, alternative mechanisms, and amino acids involved.
FMA	obo.fma	The Foundational Model of Anatomy Ontology (FMA) is a biomedical informatics ontology. It is concerned with the representation of classes or types and relationships necessary for the symbolic representation of the

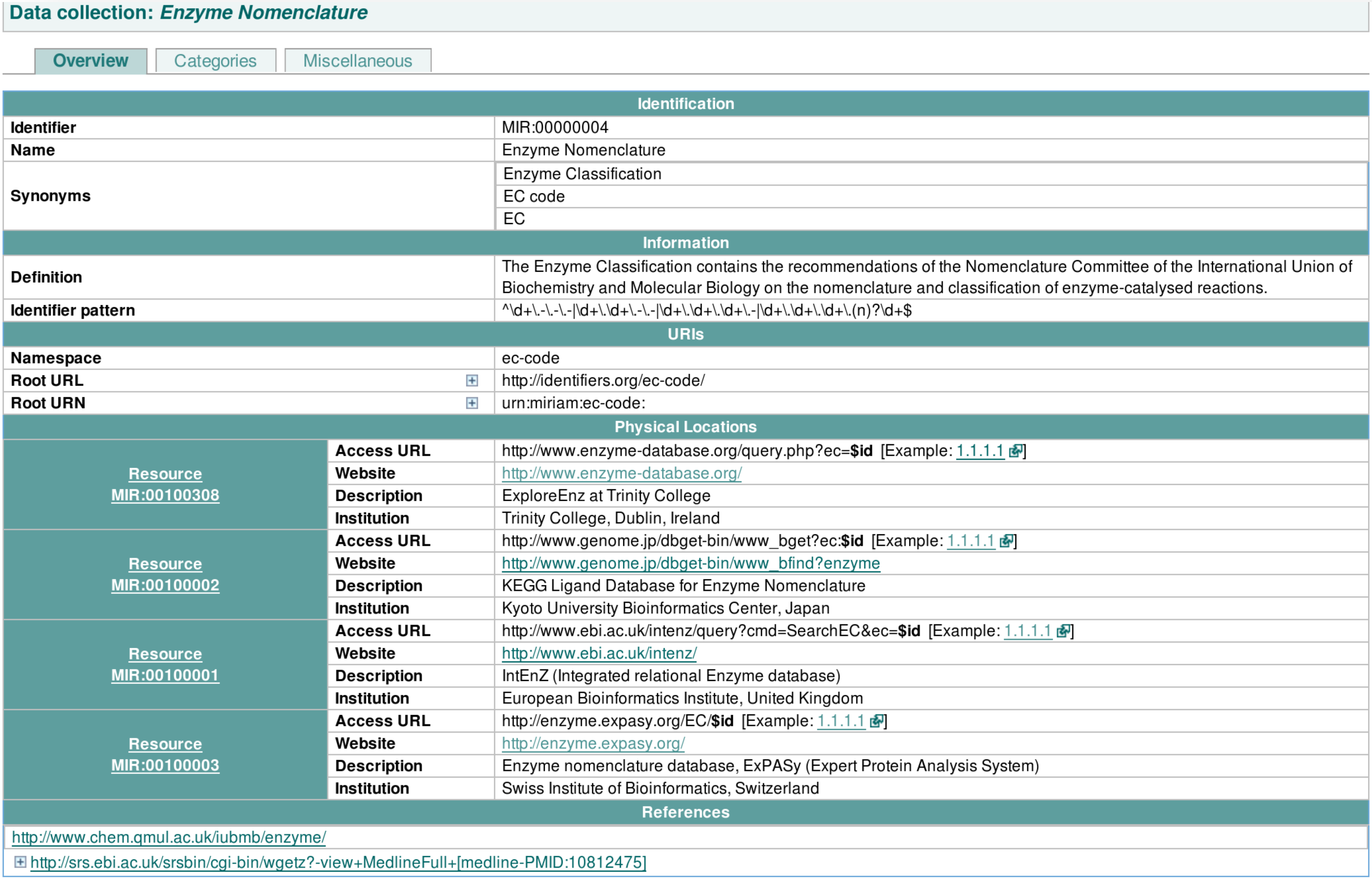
… and their associated *namespace*



Recently updated | A | B | C | D | E | F | G | H | I | J | K | L | M | N | O | P | Q | R | S | T | U | V | W | X | Y | Z

Name	Namespace	Definition
Protein Model Database	pmdb	The Protein Model DataBase (PMDB), is a database that collects manually built three dimensional protein models, obtained by different structure prediction techniques.
MIRIAM Registry collection	miriam.collection	MIRIAM Registry is an online resource created to catalogue colelctions (Gene Ontology, Taxonomy or PubMed are some examples), their URIs and the corresponding resources (physical locations), whether these are controlled vocabularies or databases.
TEDDY	biomodels.teddy	The Terminology for Description of Dynamics (TEDDY) is an ontology for dynamical behaviours, observable dynamical phenomena, and control elements of bio-models and biological systems in Systems Biology and Synthetic Biology.
UniProt	uniprot	UniProt (Universal Protein Resource) is the world's most comprehensive catalog of information on proteins. It is a central repository of protein sequence and function created by joining the information contained in Swiss-Prot, TrEMBL, and PIR.
PubChem-substance	pubchem.substance	PubChem provides information on the biological activities of small molecules. It is a component of NIH's Molecular Libraries Roadmap Initiative. PubChem Substance archives chemical substance records.
EDAM Ontology	edam	EDAM is an ontology of general bioinformatics concepts, including topics, data types, formats, identifiers and operations. EDAM provides a controlled vocabulary for the description, in semantic terms, of things such as: web services (e.g. WSDL files), applications, tool collections and packages, work-benches and workflow software, databases and ontologies, XSD data schema and data objects, data syntax and file formats, web portals and pages, resource catalogues and documents (such as scientific publications).
Ontology for Biomedical Investigations	obo.obi	The Ontology for Biomedical Investigations (OBI) project is developing an integrated ontology for the description of biological and clinical investigations. The ontology will represent the design of an investigation, the protocols and instrumentation used, the material used, the data generated and the type analysis performed on it. Currently OBI is being built under the Basic Formal Ontology (BFO).
MACiE	macie	MACiE (Mechanism, Annotation and Classification in Enzymes) is a database of enzyme reaction mechanisms. Each entry in MACiE consists of an overall reaction describing the chemical compounds involved, as well as the species name in which the reaction occurs. The individual reaction stages for each overall reaction are listed with mechanisms, alternative mechanisms, and amino acids involved.
FMA	obo.fma	The Foundational Model of Anatomy Ontology (FMA) is a biomedical informatics ontology. It is concerned with the representation of classes or types and relationships necessary for the symbolic representation of the

Data collection



Data collection: *Enzyme Nomenclature*

Identification	
Identifier	MIR:00000004
Name	Enzyme Nomenclature
Synonyms	Enzyme Classification EC code EC

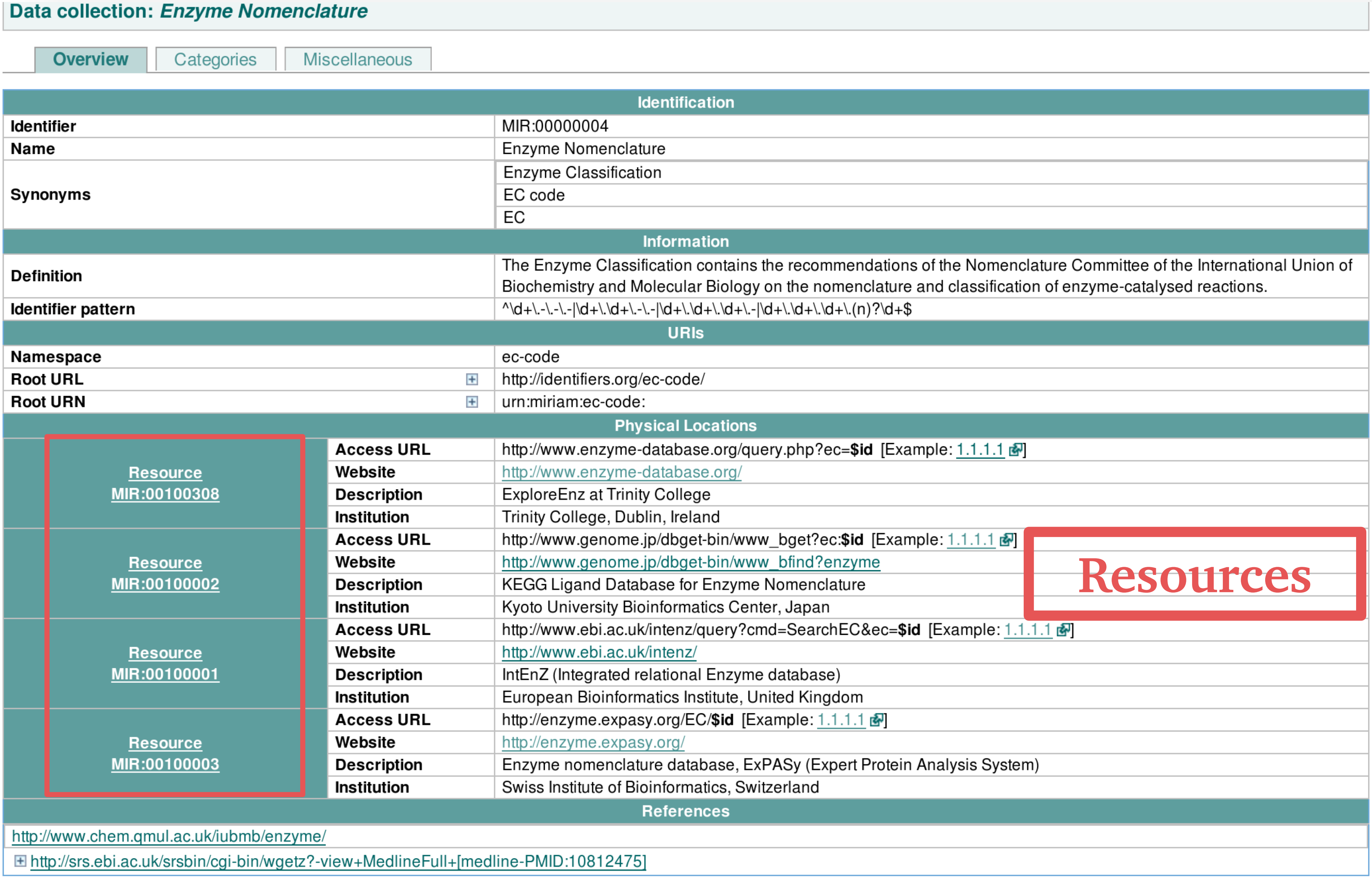
Information	
Definition	The Enzyme Classification contains the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzyme-catalysed reactions.
Identifier pattern	^\d+\.-\.-\.-|\d+\.\d+\.-\.-|\d+\.\d+\.\d+\.-|\d+\.\d+\.\d+\.(n)?\d+$

URIs	
Namespace	ec-code
Root URL ⊞	http://identifiers.org/ec-code/
Root URN ⊞	urn:miriam:ec-code:

Physical Locations		
Resource MIR:00100308	**Access URL**	http://www.enzyme-database.org/query.php?ec=**$id** [Example: 1.1.1.1]
	Website	http://www.enzyme-database.org/
	Description	ExploreEnz at Trinity College
	Institution	Trinity College, Dublin, Ireland
Resource MIR:00100002	**Access URL**	http://www.genome.jp/dbget-bin/www_bget?ec:**$id** [Example: 1.1.1.1]
	Website	http://www.genome.jp/dbget-bin/www_bfind?enzyme
	Description	KEGG Ligand Database for Enzyme Nomenclature
	Institution	Kyoto University Bioinformatics Center, Japan
Resource MIR:00100001	**Access URL**	http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&ec=**$id** [Example: 1.1.1.1]
	Website	http://www.ebi.ac.uk/intenz/
	Description	IntEnZ (Integrated relational Enzyme database)
	Institution	European Bioinformatics Institute, United Kingdom
Resource MIR:00100003	**Access URL**	http://enzyme.expasy.org/EC/**$id** [Example: 1.1.1.1]
	Website	http://enzyme.expasy.org/
	Description	Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)
	Institution	Swiss Institute of Bioinformatics, Switzerland

References
http://www.chem.qmul.ac.uk/iubmb/enzyme/
⊞ http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-view+MedlineFull+[medline-PMID:10812475]

… and associated resources



Data collection: *Enzyme Nomenclature*

Identification	
Identifier	MIR:00000004
Name	Enzyme Nomenclature
Synonyms	Enzyme Classification EC code EC

Information	
Definition	The Enzyme Classification contains the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzyme-catalysed reactions.
Identifier pattern	^\d+\.-\.-\.-|\d+\.\d+\.-\.-|\d+\.\d+\.\d+\.-|\d+\.\d+\.\d+\.(n)?\d+$

URIs	
Namespace	ec-code
Root URL ⊞	http://identifiers.org/ec-code/
Root URN ⊞	urn:miriam:ec-code:

Physical Locations		
Resource MIR:00100308	**Access URL**	http://www.enzyme-database.org/query.php?ec=**$id** [Example: 1.1.1.1]
	Website	http://www.enzyme-database.org/
	Description	ExploreEnz at Trinity College
	Institution	Trinity College, Dublin, Ireland
Resource MIR:00100002	**Access URL**	http://www.genome.jp/dbget-bin/www_bget?ec:**$id** [Example: 1.1.1.1]
	Website	http://www.genome.jp/dbget-bin/www_bfind?enzyme
	Description	KEGG Ligand Database for Enzyme Nomenclature
	Institution	Kyoto University Bioinformatics Center, Japan
Resource MIR:00100001	**Access URL**	http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&ec=**$id** [Example: 1.1.1.1]
	Website	http://www.ebi.ac.uk/intenz/
	Description	IntEnZ (Integrated relational Enzyme database)
	Institution	European Bioinformatics Institute, United Kingdom
Resource MIR:00100003	**Access URL**	http://enzyme.expasy.org/EC/**$id** [Example: 1.1.1.1]
	Website	http://enzyme.expasy.org/
	Description	Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)
	Institution	Swiss Institute of Bioinformatics, Switzerland

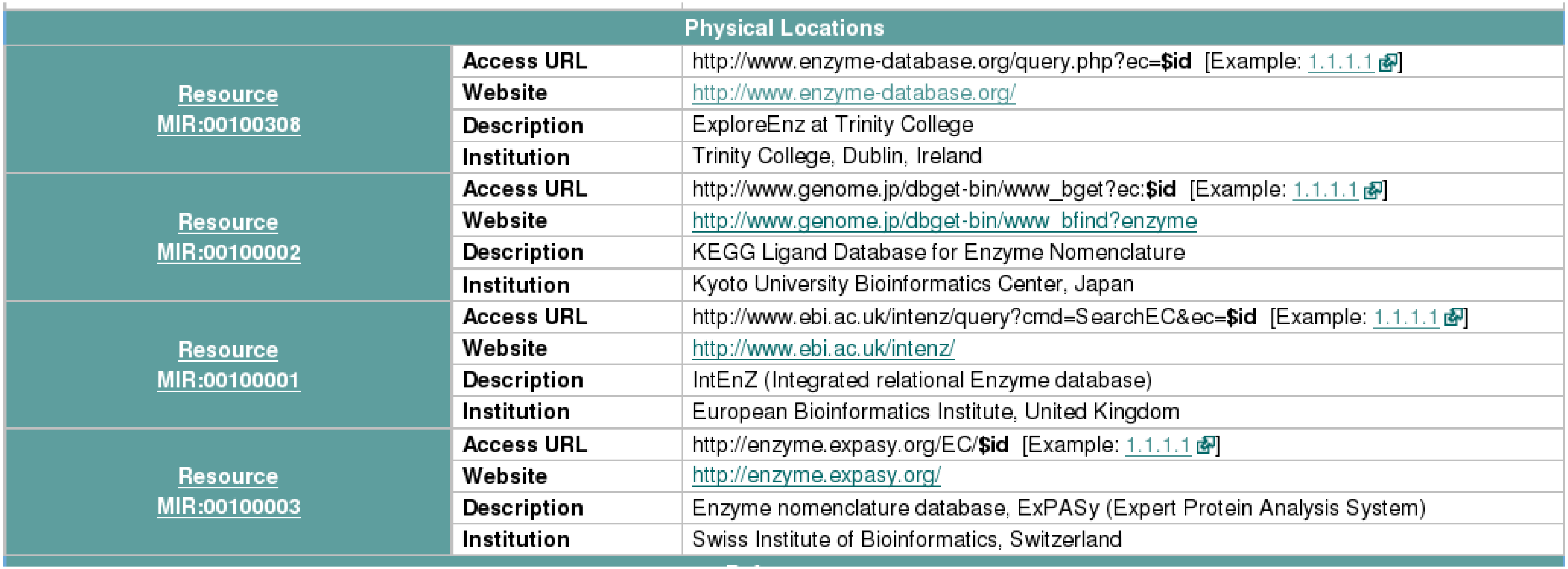
References

http://www.chem.qmul.ac.uk/iubmb/enzyme/

⊞ http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-view+MedlineFull+[medline-PMID:10812475]

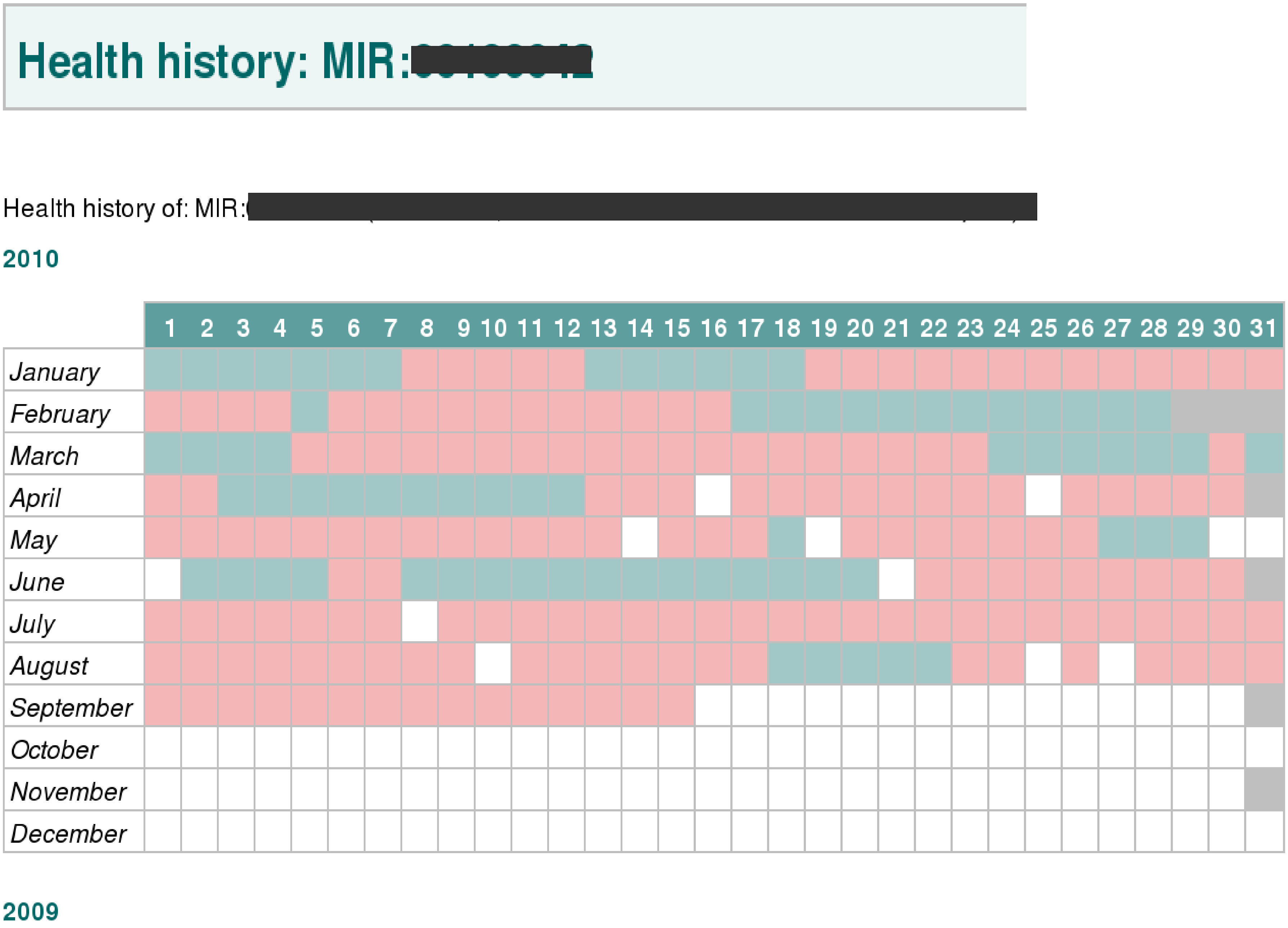
Resources

… and associated resources



Physical Locations		
Resource MIR:00100308	**Access URL**	http://www.enzyme-database.org/query.php?ec=**$id** [Example: 1.1.1.1]
	Website	http://www.enzyme-database.org/
	Description	ExploreEnz at Trinity College
	Institution	Trinity College, Dublin, Ireland
Resource MIR:00100002	**Access URL**	http://www.genome.jp/dbget-bin/www_bget?ec:**$id** [Example: 1.1.1.1]
	Website	http://www.genome.jp/dbget-bin/www_bfind?enzyme
	Description	KEGG Ligand Database for Enzyme Nomenclature
	Institution	Kyoto University Bioinformatics Center, Japan
Resource MIR:00100001	**Access URL**	http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&ec=**$id** [Example: 1.1.1.1]
	Website	http://www.ebi.ac.uk/intenz/
	Description	IntEnZ (Integrated relational Enzyme database)
	Institution	European Bioinformatics Institute, United Kingdom
Resource MIR:00100003	**Access URL**	http://enzyme.expasy.org/EC/**$id** [Example: 1.1.1.1]
	Website	http://enzyme.expasy.org/
	Description	Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)
	Institution	Swiss Institute of Bioinformatics, Switzerland

Resource reliability



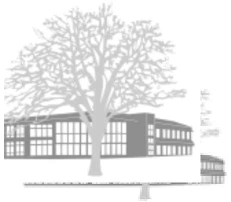
Health history: MIR:[redacted]

Health history of: MIR:[redacted]

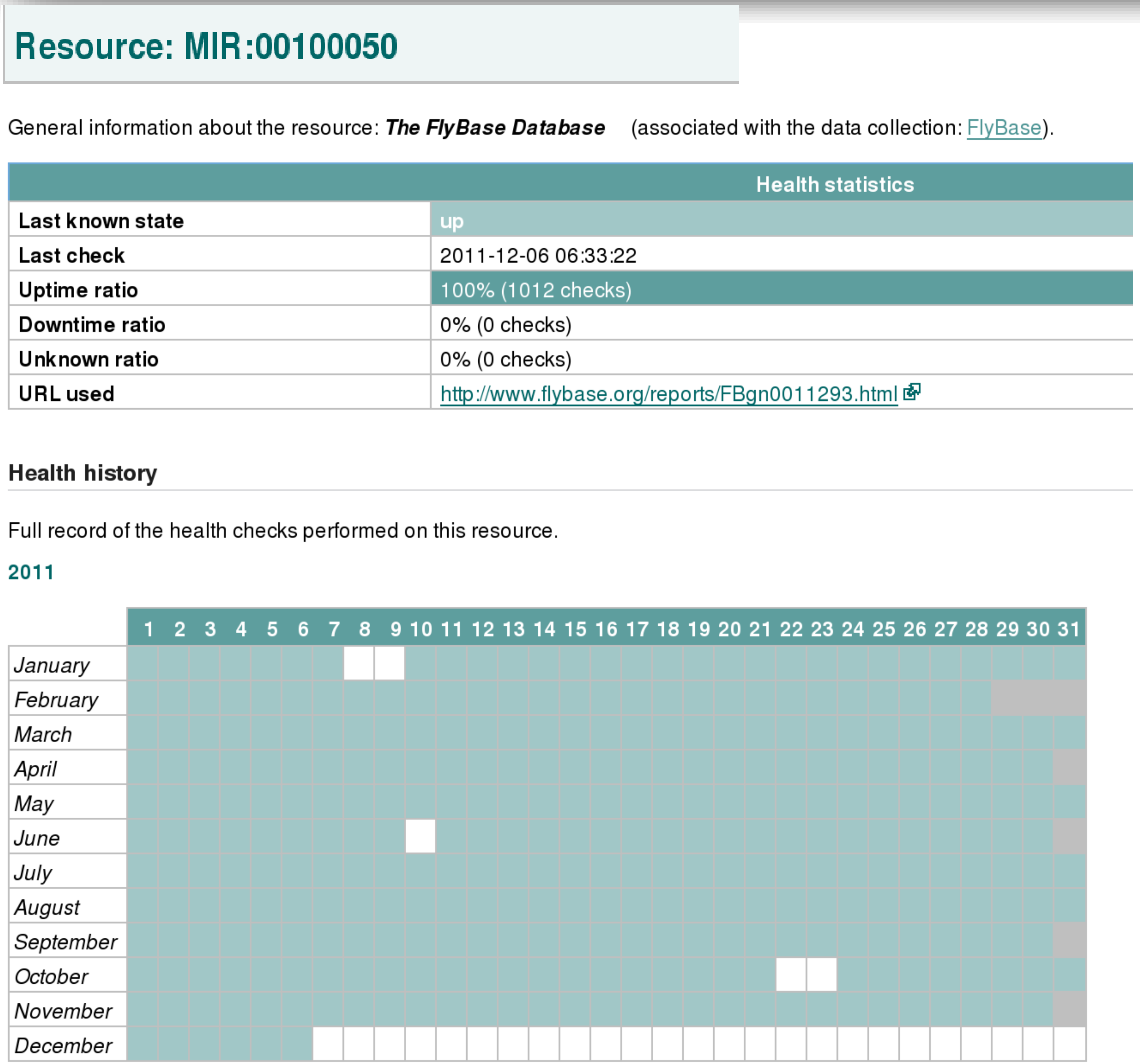
2010

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
January	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
February	🟥	🟥	🟥	🟥	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	▨	▨	▨
March	🟩	🟩	🟩	🟩	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟩
April	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟥	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥		🟥	🟥	🟥	🟥	🟥	▨
May	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥		🟥	🟥	🟥	🟩		🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩		
June		🟩	🟩	🟩	🟩	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩		🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	▨
July	🟥	🟥	🟥	🟥	🟥	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥
August	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥		🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟩	🟩	🟩	🟩	🟩	🟥	🟥		🟥		🟥	🟥	🟥	🟥
September	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥	🟥																▨
October																															
November																															▨
December																															

2009



Resource reliability



Resource: MIR:00100050

General information about the resource: ***The FlyBase Database*** (associated with the data collection: FlyBase).

Health statistics	
Last known state	up
Last check	2011-12-06 06:33:22
Uptime ratio	100% (1012 checks)
Downtime ratio	0% (0 checks)
Unknown ratio	0% (0 checks)
URL used	http://www.flybase.org/reports/FBgn0011293.html

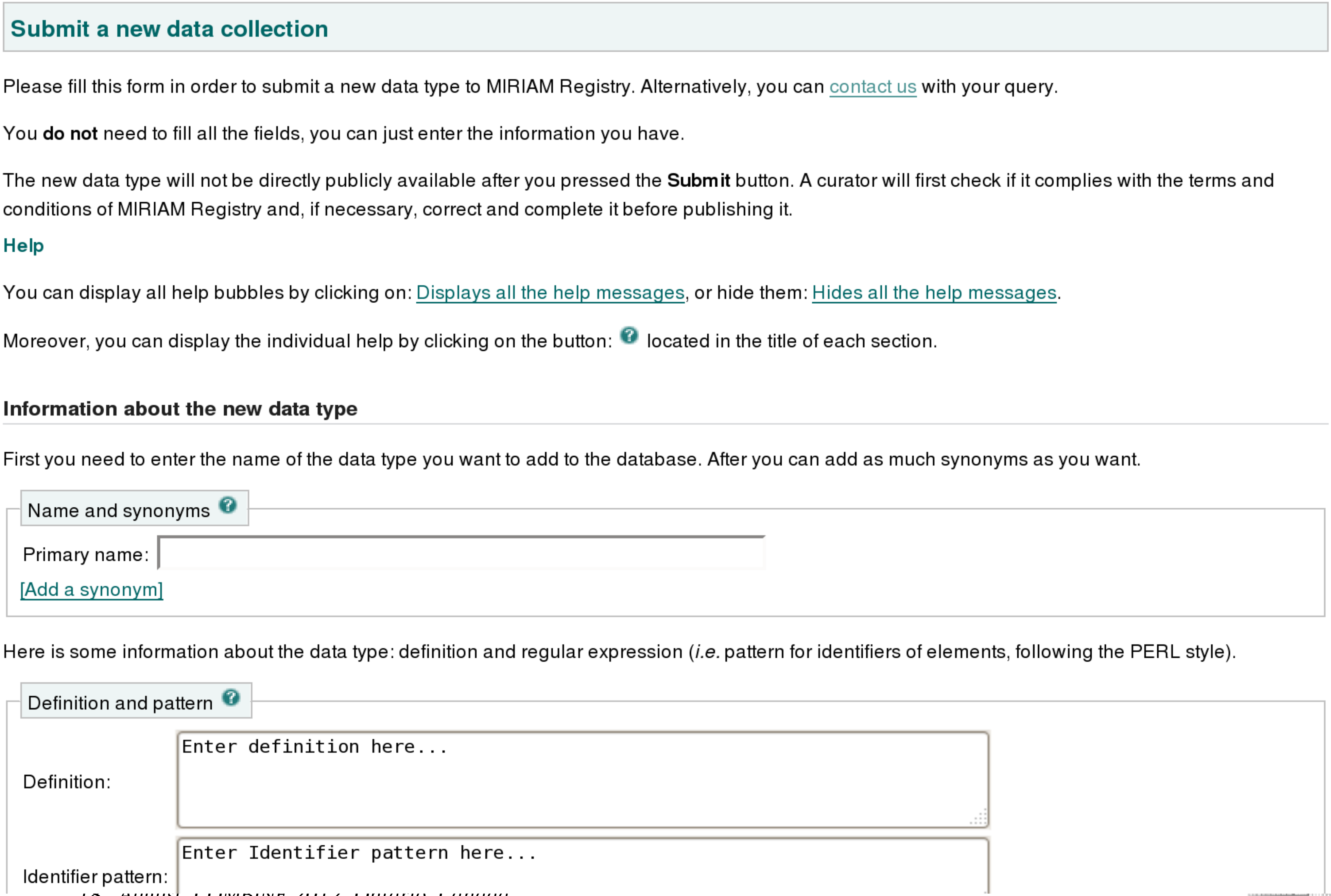
Health history

Full record of the health checks performed on this resource.

2011

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31
January	🟩	🟩	🟩	🟩	🟩	🟩	🟩	☐	☐	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
February	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	▨	▨	▨
March	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
April	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	▨
May	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
June	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	☐	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	▨
July	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
August	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
September	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	▨
October	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	☐	☐	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩
November	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	🟩	▨
December	🟩	🟩	🟩	🟩	🟩	🟩	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐	☐

New Registry submissions



Submit a new data collection

Please fill this form in order to submit a new data type to MIRIAM Registry. Alternatively, you can contact us with your query.

You **do not** need to fill all the fields, you can just enter the information you have.

The new data type will not be directly publicly available after you pressed the **Submit** button. A curator will first check if it complies with the terms and conditions of MIRIAM Registry and, if necessary, correct and complete it before publishing it.

Help

You can display all help bubbles by clicking on: Displays all the help messages, or hide them: Hides all the help messages.

Moreover, you can display the individual help by clicking on the button: ? located in the title of each section.

Information about the new data type

First you need to enter the name of the data type you want to add to the database. After you can add as much synonyms as you want.

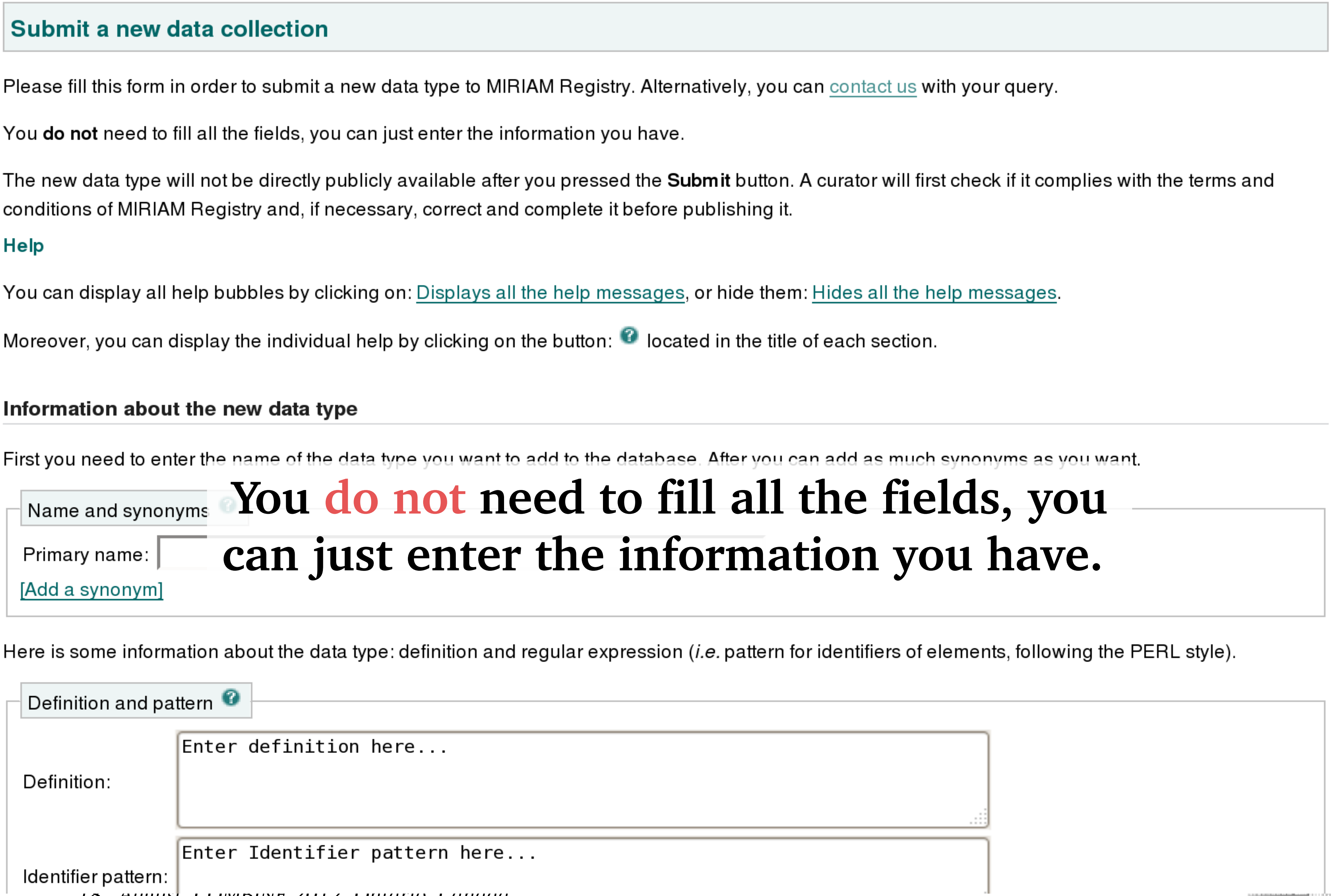
Primary name:

Here is some information about the data type: definition and regular expression (*i.e.* pattern for identifiers of elements, following the PERL style).

Definition: Enter definition here...

Identifier pattern: Enter Identifier pattern here...

New Registry submissions



Submit a new data collection

Please fill this form in order to submit a new data type to MIRIAM Registry. Alternatively, you can contact us with your query.

You **do not** need to fill all the fields, you can just enter the information you have.

The new data type will not be directly publicly available after you pressed the **Submit** button. A curator will first check if it complies with the terms and conditions of MIRIAM Registry and, if necessary, correct and complete it before publishing it.

Help

You can display all help bubbles by clicking on: Displays all the help messages, or hide them: Hides all the help messages.

Moreover, you can display the individual help by clicking on the button: ? located in the title of each section.

Information about the new data type

First you need to enter the name of the data type you want to add to the database. After you can add as much synonyms as you want.

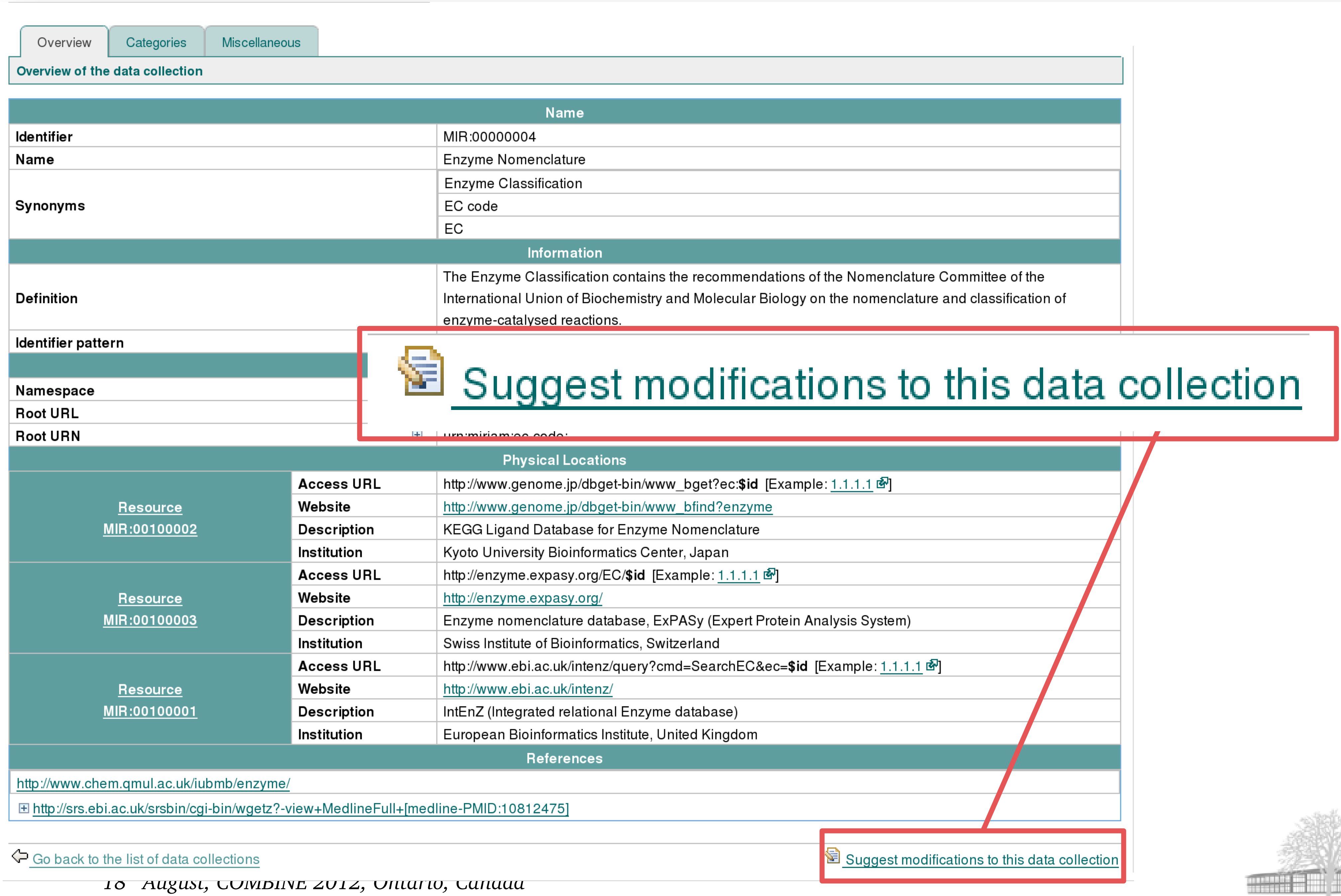
Here is some information about the data type: definition and regular expression (*i.e.* pattern for identifiers of elements, following the PERL style).

Enter definition here...

Enter Identifier pattern here...

You do not need to fill all the fields, you can just enter the information you have.

Registry: suggest changes



Overview of the data collection

Name	
Identifier	MIR:00000004
Name	Enzyme Nomenclature
Synonyms	Enzyme Classification EC code EC
Information	
Definition	The Enzyme Classification contains the recommendations of the Nomenclature Committee of the International Union of Biochemistry and Molecular Biology on the nomenclature and classification of enzyme-catalysed reactions.
Identifier pattern	
Namespace	
Root URL	
Root URN	urn:miriam:ec-code:

Physical Locations		
Resource MIR:00100002	**Access URL**	http://www.genome.jp/dbget-bin/www_bget?ec:**$id** [Example: 1.1.1.1]
	Website	http://www.genome.jp/dbget-bin/www_bfind?enzyme
	Description	KEGG Ligand Database for Enzyme Nomenclature
	Institution	Kyoto University Bioinformatics Center, Japan
Resource MIR:00100003	**Access URL**	http://enzyme.expasy.org/EC/**$id** [Example: 1.1.1.1]
	Website	http://enzyme.expasy.org/
	Description	Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)
	Institution	Swiss Institute of Bioinformatics, Switzerland
Resource MIR:00100001	**Access URL**	http://www.ebi.ac.uk/intenz/query?cmd=SearchEC&ec=**$id** [Example: 1.1.1.1]
	Website	http://www.ebi.ac.uk/intenz/
	Description	IntEnZ (Integrated relational Enzyme database)
	Institution	European Bioinformatics Institute, United Kingdom

References

http://www.chem.qmul.ac.uk/iubmb/enzyme/

⊞ http://srs.ebi.ac.uk/srsbin/cgi-bin/wgetz?-view+MedlineFull+[medline-PMID:10812475]

Go back to the list of data collections

Suggest modifications to this data collection

Suggest modifications to this data collection

Registry: summary

list of **data collections**

web services and **XML export** available

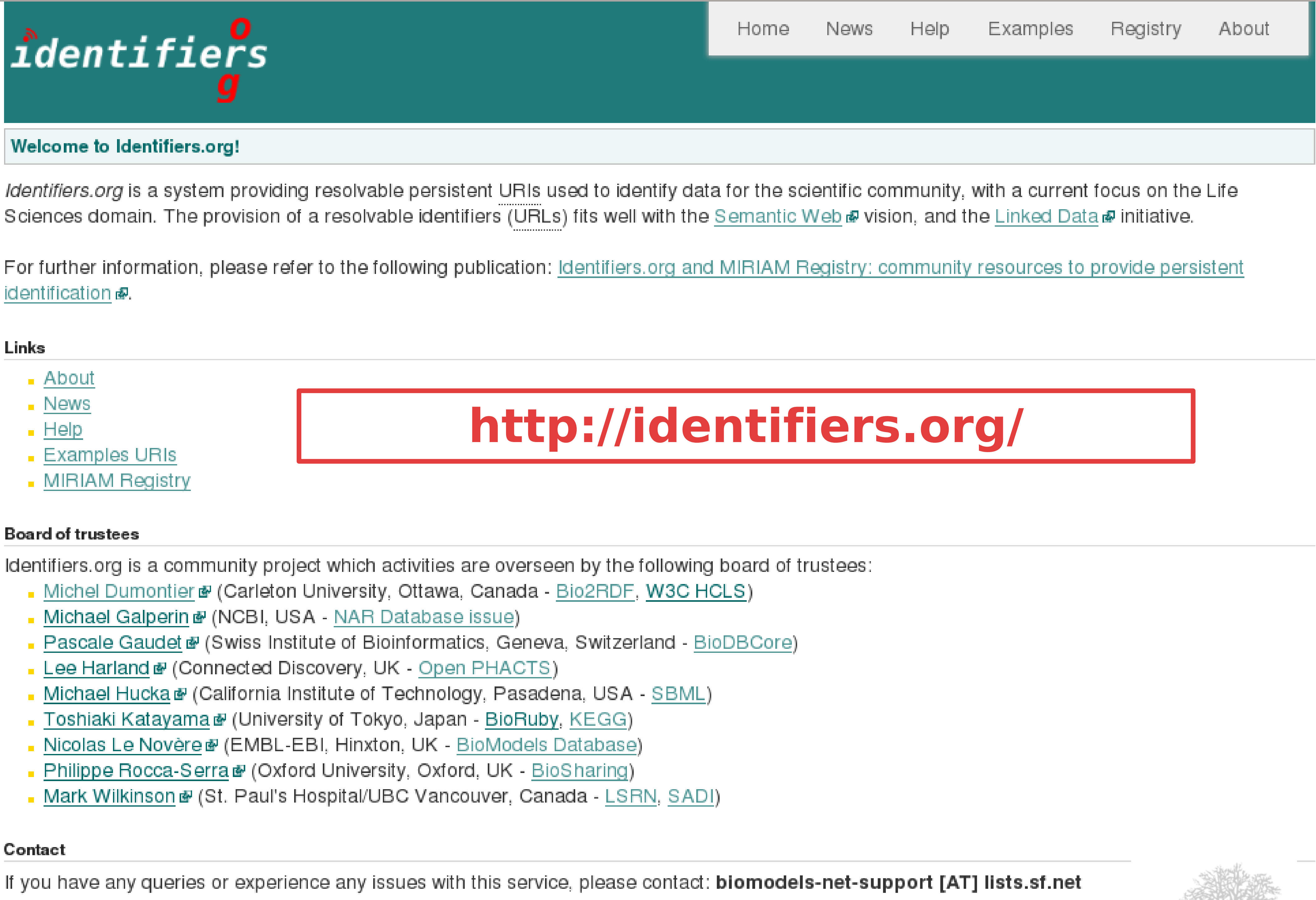
community driven (everybody can submit new data collections or suggest modification of existing records)

curated resource

systems in place to **monitor** registered web resources

unrestricted scope (currently mainly focused on Life Sciences, but the scope is potentially unlimited)

free to use



Welcome to Identifiers.org!

Identifiers.org is a system providing resolvable persistent URIs used to identify data for the scientific community, with a current focus on the Life Sciences domain. The provision of a resolvable identifiers (URLs) fits well with the Semantic Web vision, and the Linked Data initiative.

For further information, please refer to the following publication: Identifiers.org and MIRIAM Registry: community resources to provide persistent identification.

Links

- About
- News
- Help
- Examples URIs
- MIRIAM Registry

http://identifiers.org/

Board of trustees

Identifiers.org is a community project which activities are overseen by the following board of trustees:

Contact

If you have any queries or experience any issues with this service, please contact: **biomodels-net-support [AT] lists.sf.net**

Identifiers.org: basic identification

Data Collection:

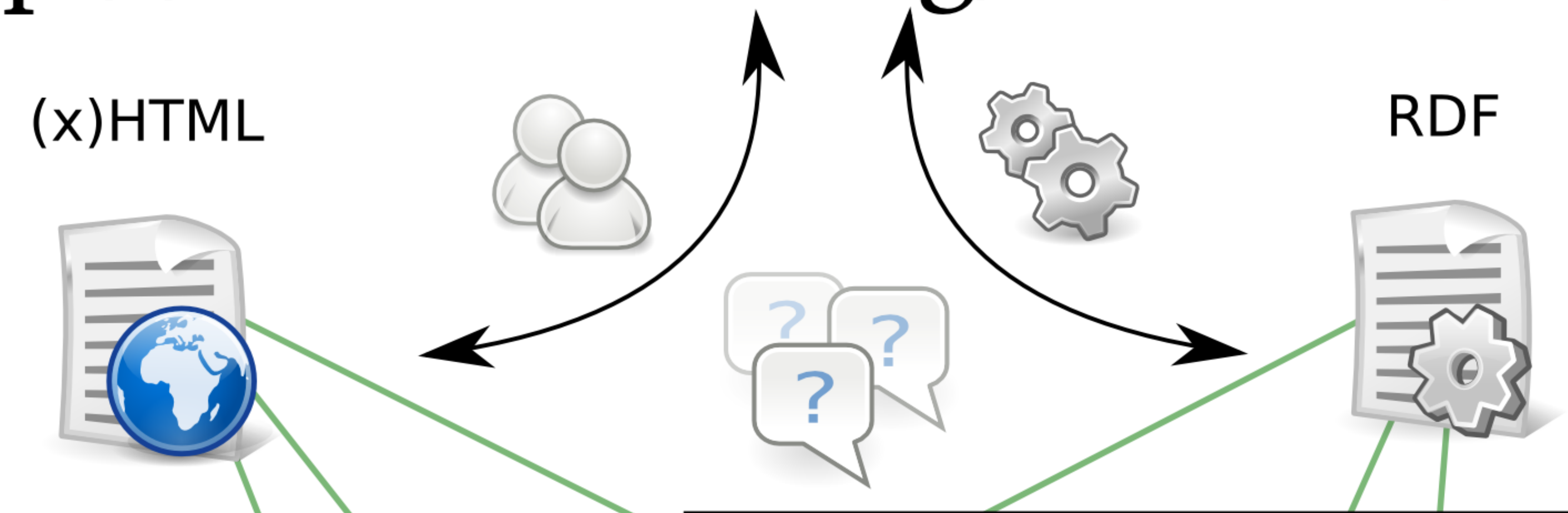
http://identifiers.org/ec-code/

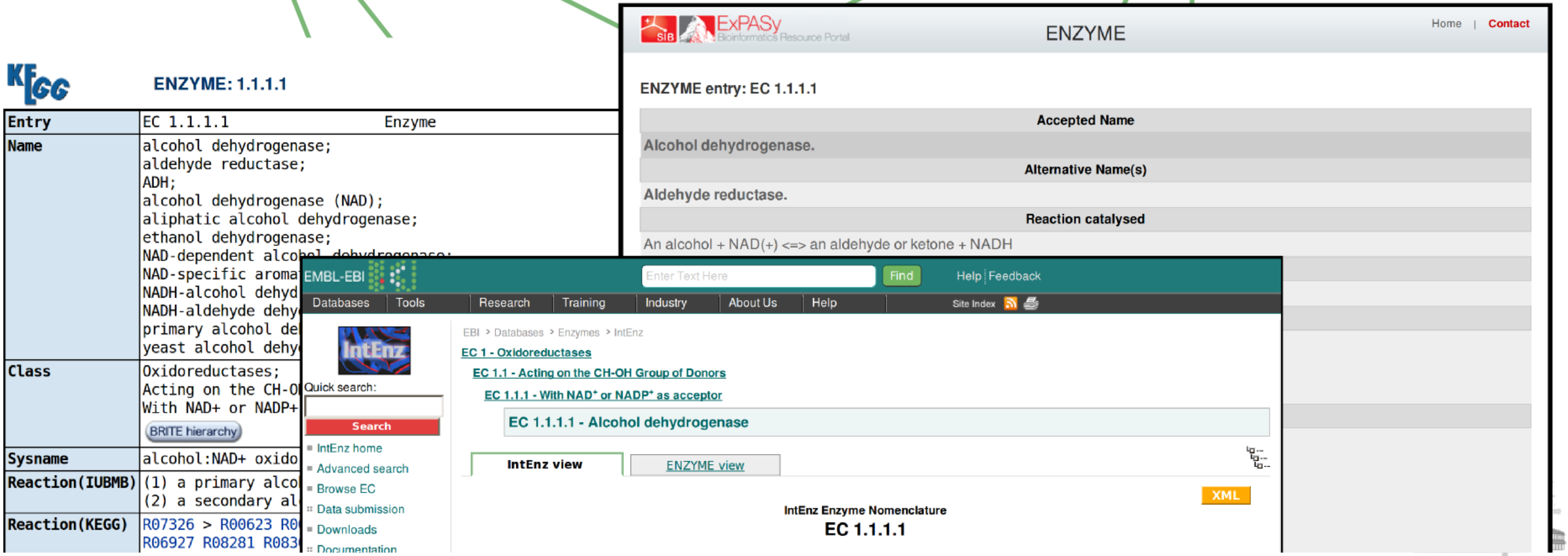
Record (location independent):

http://identifiers.org/ec-code/1.1.1.1

Identifiers.org: resolving

http://identifiers.org/ec-code/1.1.1.1





ENZYME: 1.1.1.1

Entry	EC 1.1.1.1 Enzyme
Name	alcohol dehydrogenase; aldehyde reductase; ADH; alcohol dehydrogenase (NAD); aliphatic alcohol dehydrogenase; ethanol dehydrogenase; NAD-dependent alcohol dehydrogenase; NAD-specific aroma NADH-alcohol dehyd NADH-aldehyde dehy primary alcohol de yeast alcohol dehy
Class	Oxidoreductases; Acting on the CH-O With NAD+ or NADP+ BRITE hierarchy
Sysname	alcohol:NAD+ oxido
Reaction(IUBMB)	(1) a primary alco (2) a secondary al
Reaction(KEGG)	R07326 > R00623 R0 R06927 R08281 R083

ENZYME

ENZYME entry: EC 1.1.1.1

Accepted Name

Alcohol dehydrogenase.

Alternative Name(s)

Aldehyde reductase.

Reaction catalysed

An alcohol + NAD(+) <=> an aldehyde or ketone + NADH

EBI > Databases > Enzymes > IntEnz

EC 1 - Oxidoreductases

EC 1.1 - Acting on the CH-OH Group of Donors

EC 1.1.1 - With NAD$^+$ or NADP$^+$ as acceptor

EC 1.1.1.1 - Alcohol dehydrogenase

IntEnz Enzyme Nomenclature

EC 1.1.1.1

Choice of resources (HTML version)

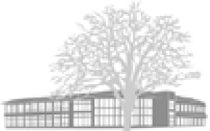
http://identifiers.org/ec-code/1.1.1.1

4 physical locations (or resources) are available for accessing *1.1.1.1* (from Enzyme Nomenclature):

KEGG Ligand Database for Enzyme Nomenclature Kyoto University Bioinformatics Center *Japan* (Uptime: 100%)	**ExploreEnz at Trinity College** Trinity College, Dublin *Ireland* (Uptime: 100%)
IntEnZ (Integrated relational Enzyme database) European Bioinformatics Institute *United Kingdom* (Uptime: 100%)	**Enzyme nomenclature database, ExPASy (Expert Protein Analysis System)** Swiss Institute of Bioinformatics *Switzerland* (Uptime: 99%)

Powered by MIRIAM Registry

Information also available in: RDF



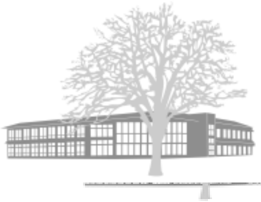
Identifiers.org: more identification

Custom queries + format:

Response format:
    http://identifiers.org/ec-code/1.1.1.1?format=rdfxml

Specified resource:
    http://identifiers.org/ec-code/1.1.1.1?resource=MIR:00100001

profiles (predefined resolution locations):
    http://identifiers.org/ec-code/1.1.1.1?profile=demo



Resolving: basic

http://identifiers.org/pubmed/16333295

http://identifiers.org/pubmed/16333295

4 physical locations (or resources) are available for accessing *16333295* (from PubMed):

HubMed Alfred D. Eaton *United Kingdom* (Uptime: 96%)	**SRS@EBI** European Bioinformatics Institute *United Kingdom* (Uptime: 100%)
NCBI PubMed National Center for Biotechnology Information *USA* (Uptime: 100%)	**CiteXplore** European Bioinformatics Institute *United Kingdom* (Uptime: 99%)

Powered by MIRIAM Registry

Information also available in: RDF

Resolving: using a specific resource

http://identifiers.org/pubmed/16333295?resource=MIR:00100032

http://identifiers.org/pubmed/16333295

Access to 16333295 (from PubMed) using the resource MIR:00100032.
Entity available from **4** providers, for more information please refer to: http://identifiers.org/pubmed/16333295.

CiteXplore - citation details

Pubmed Id	16333295
Title	**Minimum information requested in the annotation of biochemical models (MIRIAM).**
Authors	Le Novère N , Finney A , Hucka M , Bhalla US , Campagne F , Collado-Vides J , Crampin EJ , Halstead M , Klipp E , Mendes P See all ..
Affiliation	European Bioinformatics Institute, Hinxton, CB10 1SD, UK. lenov@ebi.ac.uk
Language	English
Journal	Nat. Biotechnol. (ISSN:1087-0156) (ESSN: 1546-1696) [2005 Dec; Volume: 23 (Issue: 12)] Page info: 1509-15
Publication type	Journal Article;
Full text article	doi>
XML	XML
Abstract	Most of the published quantitative models in biology are lost for the community because they are either not made available or they are insufficiently characterized to allow them to be reused. The lack of a standard description format, lack of stringent reviewing and authors' carelessness are the main causes for incomplete model descriptions. With today's increased interest in detailed biochemical models, it is necessary to define a minimum quality standard for the encoding of those models. We propose a set of rules for curating quantitative models of biological systems. These rules define procedures for encoding and annotating models represented in machine-readable form. We believe their application will enable users to (i) have confidence that curated models are an accurate reflection of their associated reference descriptions, (ii) search collections of curated models with precision, (iii) quickly identify the biological phenomena that a given curated model or model constituent represents and (iv) facilitate model reuse and composition into large subcellular models.
Cited by	Cited: 102 Times | Web of Science? Times Cited:146 Dry work in a wet world: computation in systems biology. 4 cite(s) Kahlem P, Birney E. *Mol. Syst. Biol. (2006)* ⊞ Show Context

profiles

Profiles:

list of data collections

subset of the Registry's content

each data collection in a profile can have settings (such as **one** preferred resource)

can be defined at the level of an institution, project, individual, ...

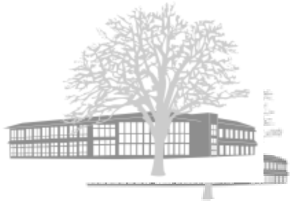
can be public (shareable) or private (protected by a key)

centrally managed in the Registry (*user interface in progress*)

have a unique shortname (used for identification purposes in URLs)

provide access to a custom XML export

Provide access to custom web services (*in progress*)



Resolving: basic

http://identifiers.org/obo.go/GO:0006915

http://identifiers.org/obo.go/GO:0006915

4 physical locations (or resources) are available for accessing *GO:0006915* (from Gene Ontology):

GO Browser
The Jackson Laboratory

USA

(Uptime: 100%)

QuickGO (Gene Ontology browser)
European Bioinformatics Institute

United Kingdom

(Uptime: 100%)

GO through BioPortal
National Center for Biomedical Ontology, Stanford

USA

(Uptime: 97%)

AmiGO
The Gene Ontology Consortium

USA

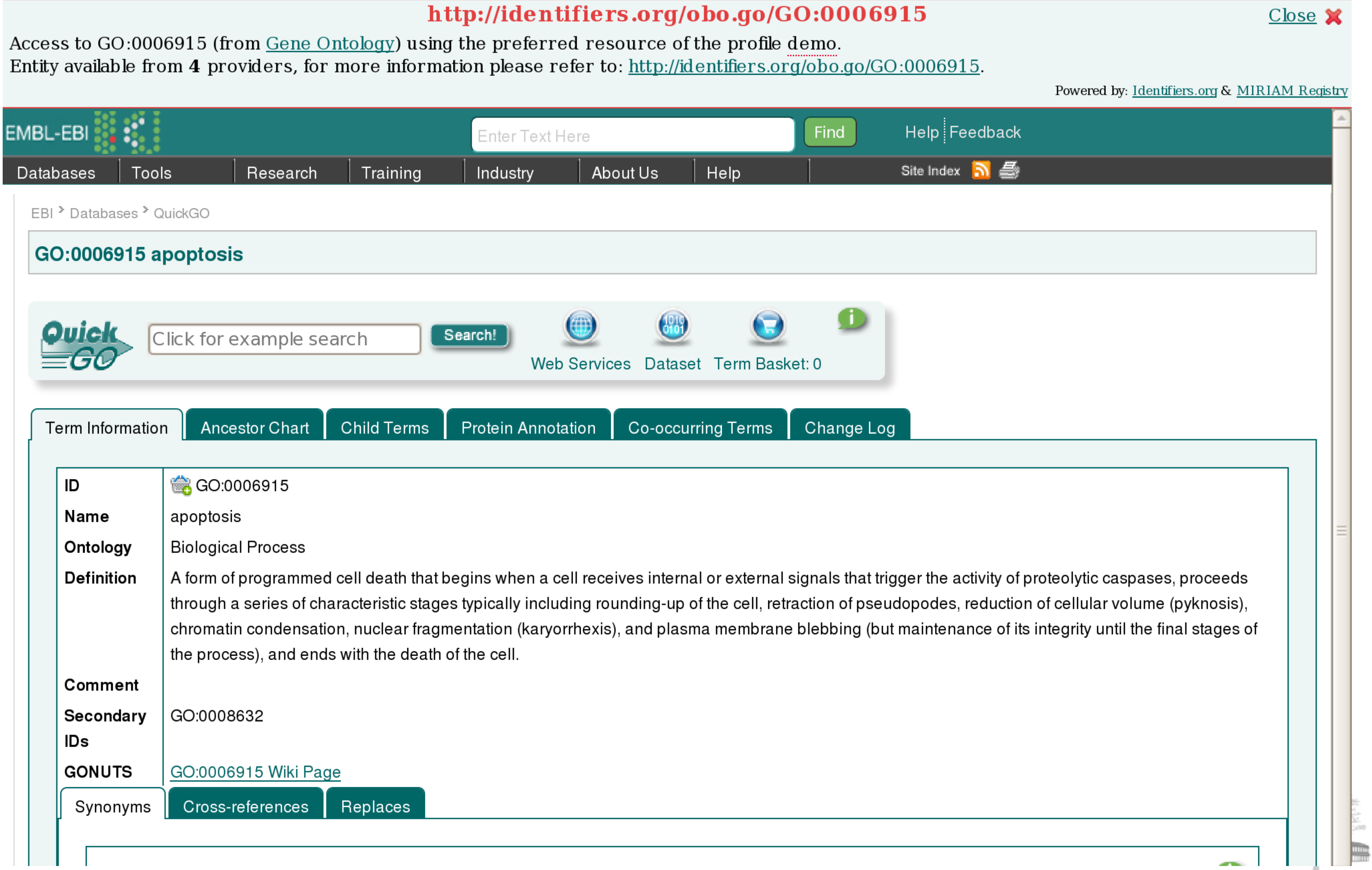
(Uptime: 99%)

Powered by MIRIAM Registry

Information also available in: RDF

Resolving using a specific profile

http://identifiers.org/obo.go/GO:0006915?profile=demo



http://identifiers.org/obo.go/GO:0006915

Close

Access to GO:0006915 (from Gene Ontology) using the preferred resource of the profile demo.
Entity available from **4** providers, for more information please refer to: http://identifiers.org/obo.go/GO:0006915.

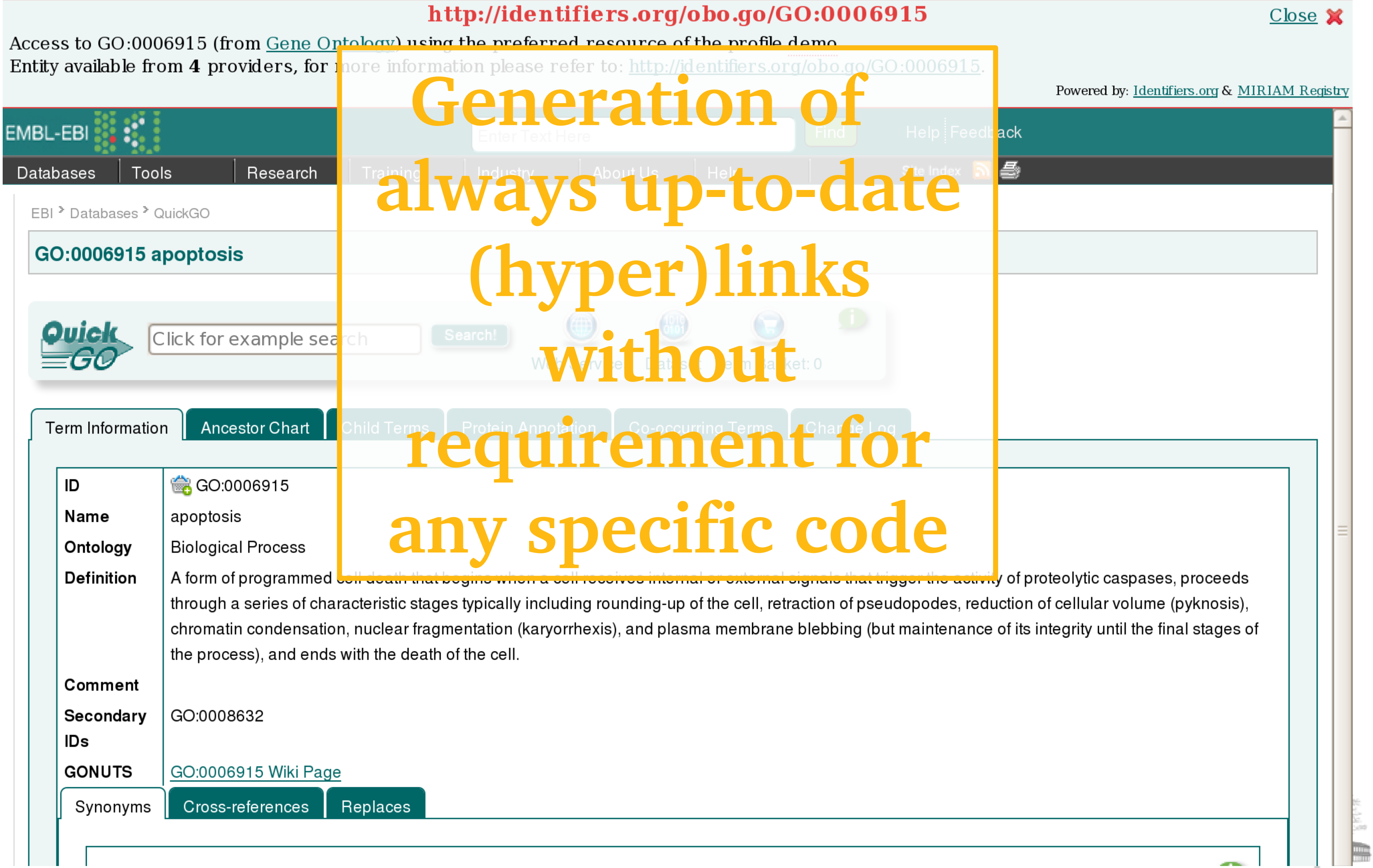
Powered by: Identifiers.org & MIRIAM Registry

GO:0006915 apoptosis

ID	GO:0006915
Name	apoptosis
Ontology	Biological Process
Definition	A form of programmed cell death that begins when a cell receives internal or external signals that trigger the activity of proteolytic caspases, proceeds through a series of characteristic stages typically including rounding-up of the cell, retraction of pseudopodes, reduction of cellular volume (pyknosis), chromatin condensation, nuclear fragmentation (karyorrhexis), and plasma membrane blebbing (but maintenance of its integrity until the final stages of the process), and ends with the death of the cell.
Comment	
Secondary IDs	GO:0008632
GONUTS	GO:0006915 Wiki Page

Resolving using a specific profile

http://identifiers.org/obo.go/GO:0006915?profile=demo



http://identifiers.org/obo.go/GO:0006915

Access to GO:0006915 (from Gene Ontology) using the preferred resource of the profile demo.
Entity available from 4 providers, for more information please refer to: http://identifiers.org/obo.go/GO:0006915.

Powered by: Identifiers.org & MIRIAM Registry

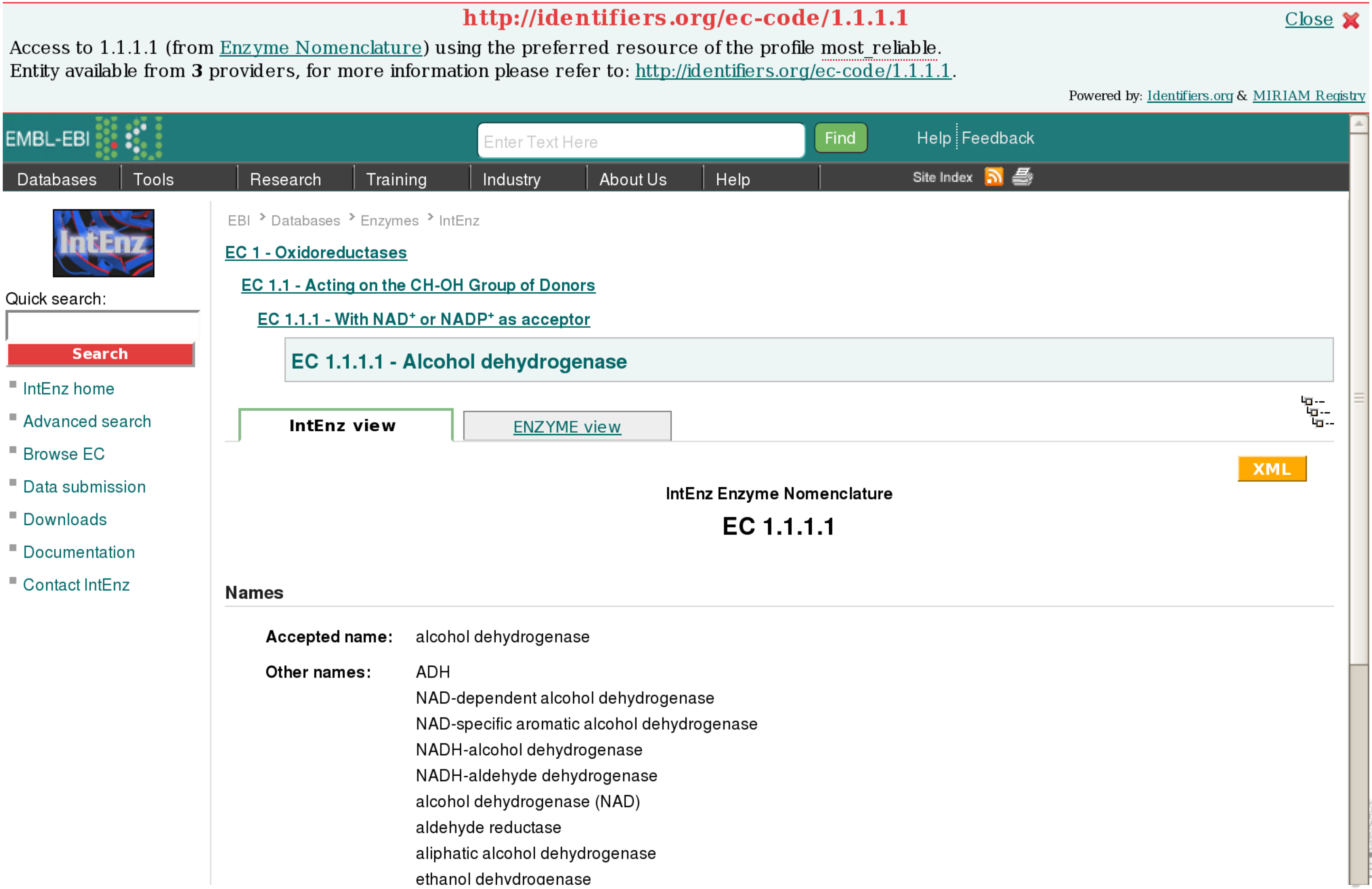
GO:0006915 apoptosis

ID	GO:0006915
Name	apoptosis
Ontology	Biological Process
Definition	A form of programmed cell death that begins when a cell receives internal or external signals that trigger the activity of proteolytic caspases, proceeds through a series of characteristic stages typically including rounding-up of the cell, retraction of pseudopodes, reduction of cellular volume (pyknosis), chromatin condensation, nuclear fragmentation (karyorrhexis), and plasma membrane blebbing (but maintenance of its integrity until the final stages of the process), and ends with the death of the cell.
Comment	
Secondary IDs	GO:0008632
GONUTS	GO:0006915 Wiki Page

Generation of always up-to-date (hyper)links without requirement for any specific code

Resolving: using the most reliable resource

http://identifiers.org/ec-code/1.1.1.1?profile=most_reliable



http://identifiers.org/ec-code/1.1.1.1

Access to 1.1.1.1 (from Enzyme Nomenclature) using the preferred resource of the profile most_reliable.
Entity available from **3** providers, for more information please refer to: http://identifiers.org/ec-code/1.1.1.1.

Powered by: Identifiers.org & MIRIAM Registry

EC 1 - Oxidoreductases

EC 1.1 - Acting on the CH-OH Group of Donors

EC 1.1.1 - With NAD$^+$ or NADP$^+$ as acceptor

EC 1.1.1.1 - Alcohol dehydrogenase

IntEnz Enzyme Nomenclature

EC 1.1.1.1

Names

Accepted name: alcohol dehydrogenase

Other names:
ADH
NAD-dependent alcohol dehydrogenase
NAD-specific aromatic alcohol dehydrogenase
NADH-alcohol dehydrogenase
NADH-aldehyde dehydrogenase
alcohol dehydrogenase (NAD)
aldehyde reductase
aliphatic alcohol dehydrogenase
ethanol dehydrogenase

Identifiers.org: short summary

provides **uniform identifiers** for Life Sciences

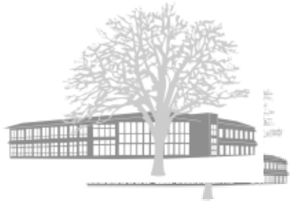
- **unambiguous**
- **perennial**
- **directly resolvable**
- standard compliant (**URIs**)
- location independent

identifiers available at **multiple levels** (data collection, resource, and data record)

customisable behaviours (formats available, preferred resource, ...)

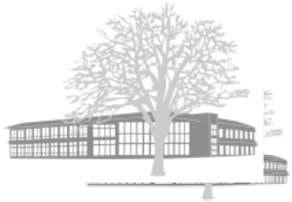
responses encoded in **HTML** and **RDF** (either requested explicitly in the URLs or via *content negotiation*)

built on the ***MIRIAM Registry***



What to store in the database?

http://identifiers.org/ec-code/1.1.1.1

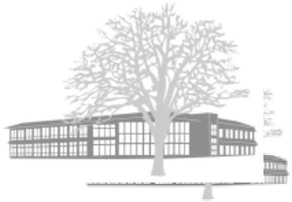


What to use in your import/export formats?

http://identifiers.org/ec-code/1.1.1.1

What to use in your user interfaces?

http://identifiers.org/ec-code/1.1.1.1



Where to get the data?

http://identifiers.org/ec-code/1.1.1.1

Acknowledgements

Camille Laibe

Identifiers.org board of trustees:

- Michel Dumontier (WC3 HCLS, Bio2RDF)
- Pascale Gaudet (BioDBCore)
- Michael Hucka (SBML)
- Nicolas Le Novère
- Philippe Rocca-Serra (BioSharing)
- Mark Wilkinson (SADI, LSRN)
- Lee Harland (Open PHACTS)
- Toshiaki Katayama (BioRuby, KEGG)
- Michael Galperin (NAR Database issue)

Biomodels-net-support
@lists.sf.net

Open PHACTS
Open Pharmacological Space

EMBL

BBSRC
bioscience for the future

NIGMS

ELIXIR

Acknowledgements

Camille Laibe

Identifiers.org board of trustees:

- Michel Dumontier (WC3 HCLS, Bio2RDF)
- Pascale Gaudet (BioDBCore)
- Michael Hucka (SBML)
- Nicolas Le Novère
- Philippe Rocca-Serra (BioSharing)
- Mark Wilkinson (SADI, LSRN)
- Lee Harland (Open PHACTS)
- Toshiaki Katayama (BioRuby, KEGG)
- Michael Galperin (NAR Database issue)

The Systems Biology community for its continued support and feedback

Biomodels-net-support
@lists.sf.net

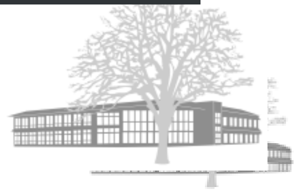
Open PHACTS
Open Pharmacological Space

EMBL

BBSRC
bioscience for the future

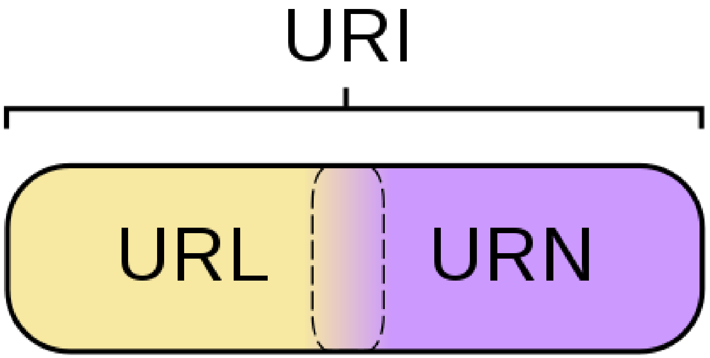
NIGMS

ELIXIR



URI: standard way to encode identifiers

A **Uniform Resource Identifier** (URI) is a string of characters used to **identify** a resource.



Uniform Resource Name (URN)

identifies a resource but does not imply its availability

Uniform Resource Locator (URL)

specifies where a resource is available on the Internet

Examples of URIs

http://www.ebi.ac.uk/

http://en.wikipedia.org/wiki/Uniform_Resource_Identifier

https://www.ebi.ac.uk/chembldb/

ftp://public.ftp-servers.example.com/mydirectory/myfile.txt

ftp://ftp.uniprot.org/pub/databases/uniprot/uniref/

urn:ietf:rfc:2648

file:///home/username/presentation.pdf

…

EMBL-EBI

MIRIAM URNs

Namespace

Identifies a data collection

Entity identifier

Identifies a data entry within the data collection

Human calmodulin: P62158 in UniProt

↳ `urn:miriam:uniprot:P62158`

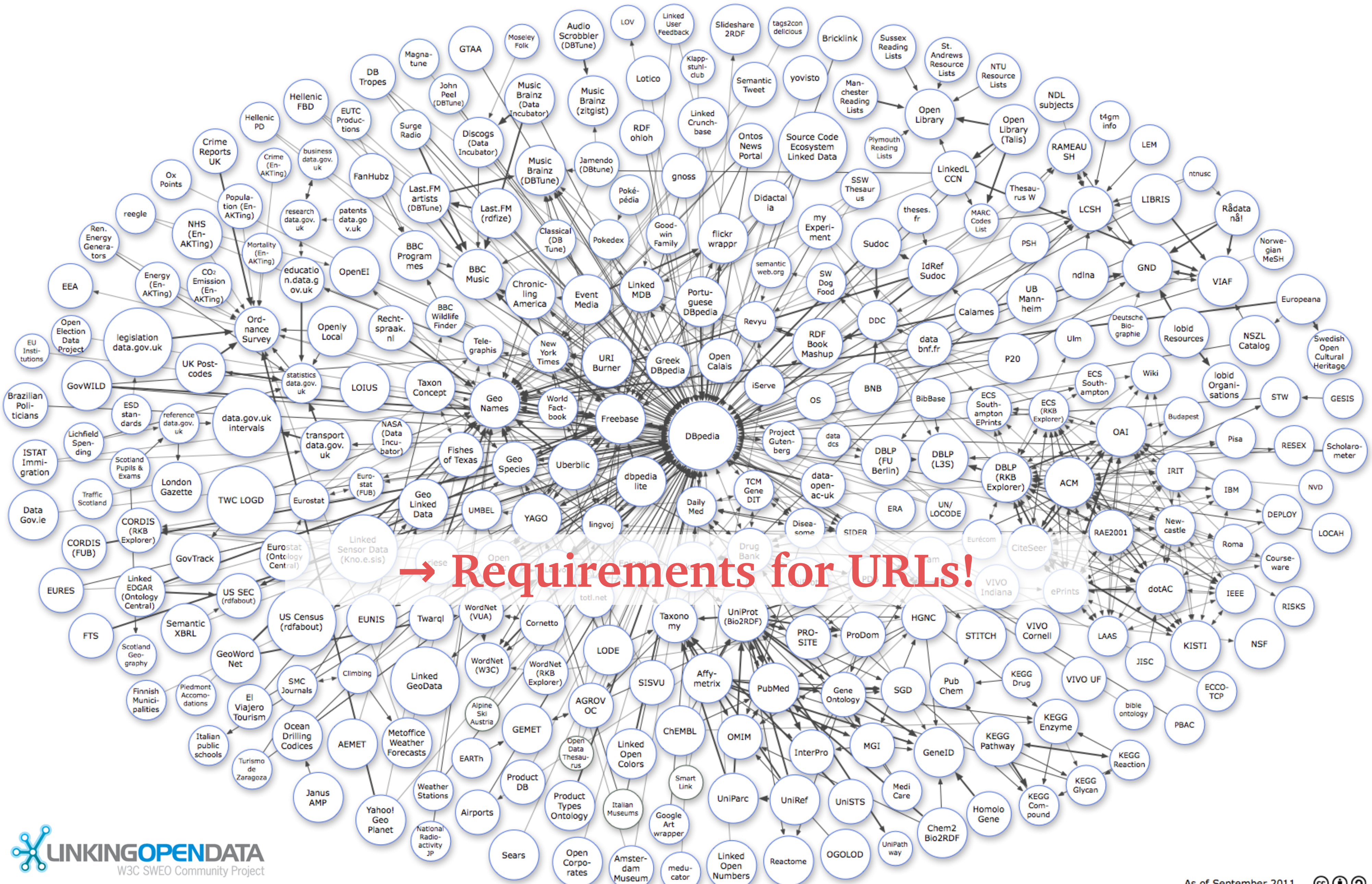
Alcohol dehydrogenase: 1.1.1.1 in EC code

↳ `urn:miriam:ec-code:1.1.1.1`

Activation of MAPKK activity: GO:0000186 in Gene Ontology

↳ `urn:miriam:obo.go:GO%3A0000186`

Linking Open Data



→ Requirements for URLs!

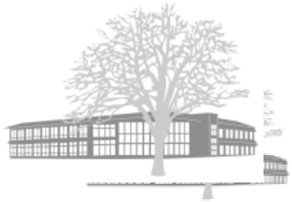
As of September 2011

Semantic Web

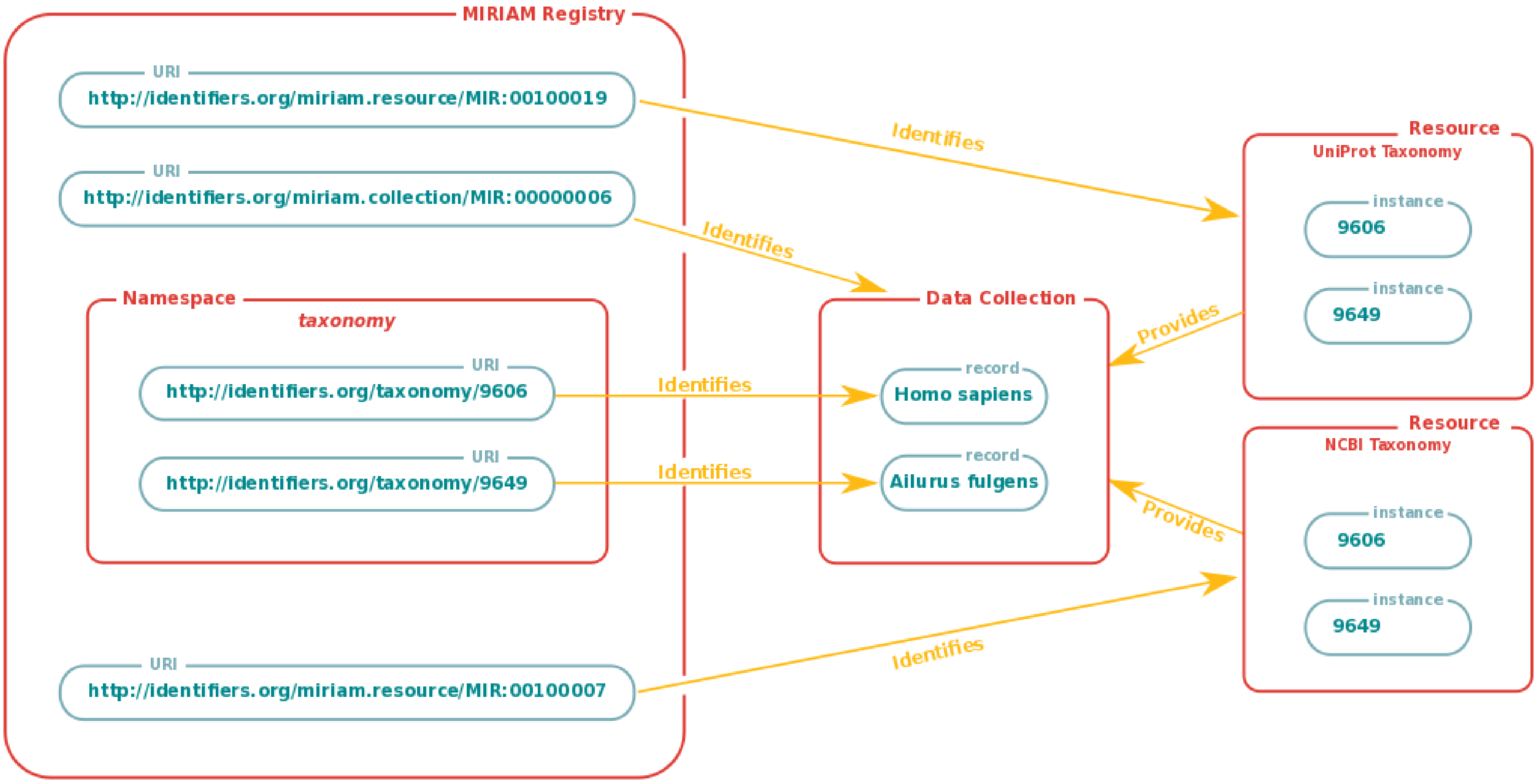
The **Semantic Web** provides a **common framework** that allows **data** to be **shared** and **reused** across application, enterprise, and community boundaries.

Goal: allows **machines** to understand the **semantics**, or meaning, of information on the **World Wide Web**.

Collaborative effort led by **W3C** with participation from a large number of researchers and industrial partners.



Nomenclature



MIRIAM Registry

URI
http://identifiers.org/miriam.resource/MIR:00100019

URI
http://identifiers.org/miriam.collection/MIR:00000006

Namespace
taxonomy

URI
http://identifiers.org/taxonomy/9606

URI
http://identifiers.org/taxonomy/9649

URI
http://identifiers.org/miriam.resource/MIR:00100007

Identifies

Identifies

Identifies

Identifies

Identifies

Data Collection

record
Homo sapiens

record
Ailurus fulgens

Provides

Provides

Resource
UniProt Taxonomy

instance
9606

instance
9649

Resource
NCBI Taxonomy

instance
9606

instance
9649

MIRIAM Registry overview

MIRIAM Database

collection catalogue

XML exports available

MIRIAM Web Services

programming interface for querying database

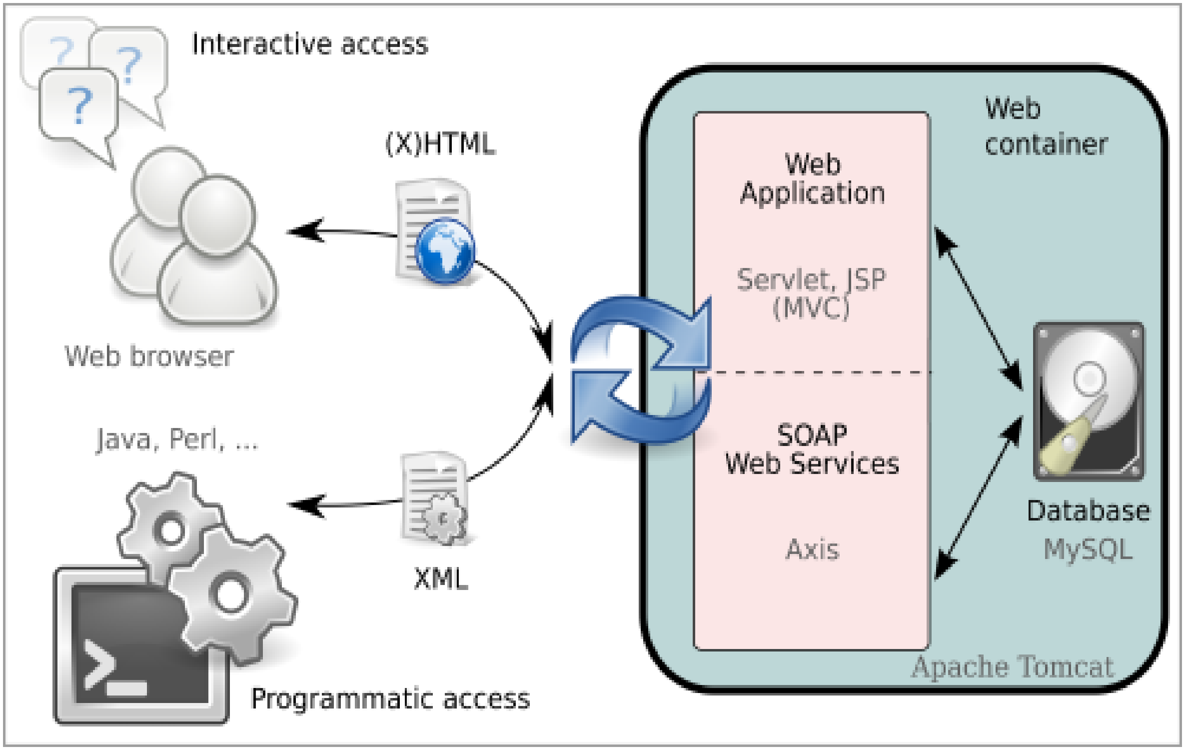
MIRIAM Library

ready to use Java code to query and access

MIRIAM Web Site

web page for browsing and querying

allows new submission (curated)



Laibe and Le Novère.
MIRIAM Resources: tools to generate and resolve robust cross-references in Systems Biology.
BMC Systems Biology, 2007