



L'API Florilège

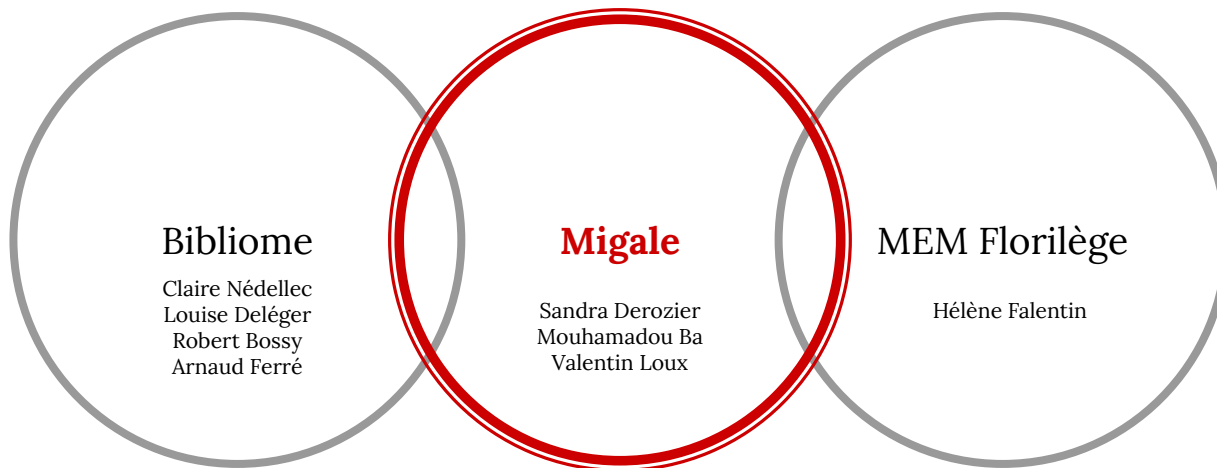


<http://migale.jouy.inra.fr/florilege-api/>

Florilege

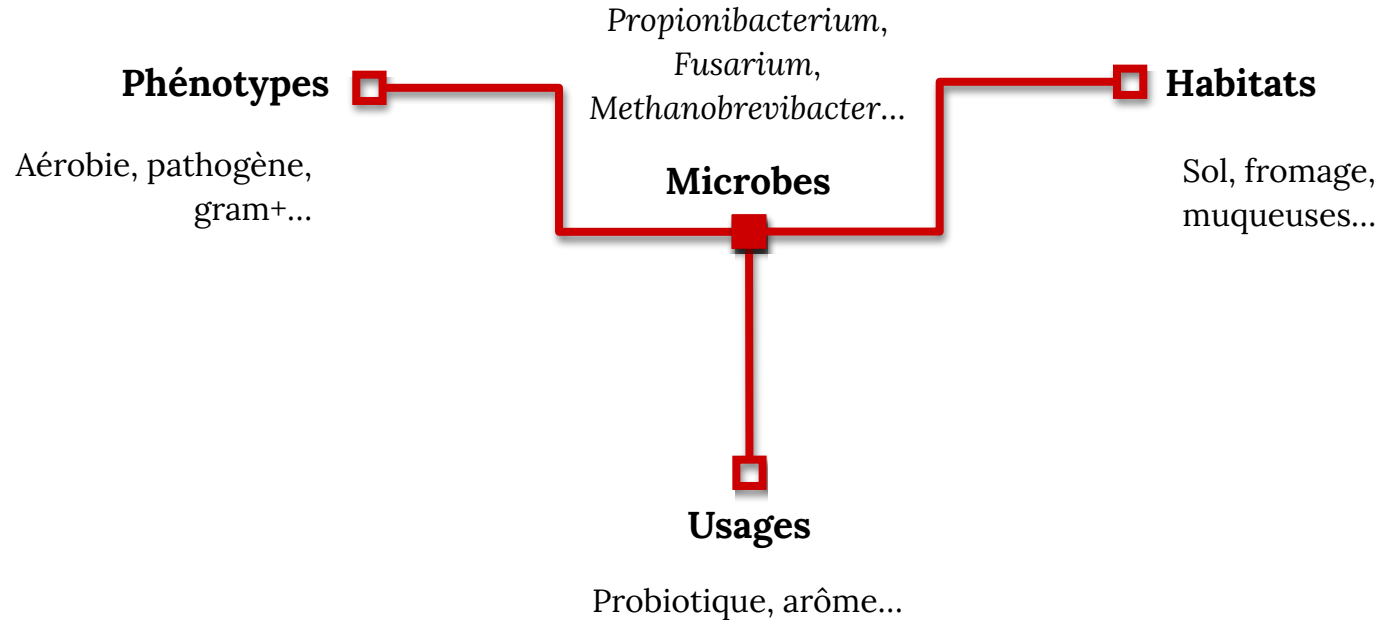


*Florilège est une base de données de référence
sur la biodiversité microbienne.*





*Florilège est une base de données de référence
sur la biodiversité microbienne.*





Florilège est une base de données de référence sur la biodiversité microbienne.

Web UI



Florilège, a database gathering microbial habitats, phenotypes and uses

- Home
- Taxon lives in Habitat
- Habitat may be inhabited by Taxon
- Taxon exhibits Phenotype
- Phenotype is exhibited by Taxon
- Taxon studied for Use
- Use involves Taxon
- Advanced search
- About Florilège
- Help

Florilège is a database of habitats, phenotypes and uses of food microbe flora

It aims to gather, in a unified representation, public information on food microbes with a focus on positive flora (microorganisms involved in transformation, bioconservation or probiotics).

Are you ready to explore the Florilège database?



- Where do a microbe or a family of microorganisms live?
 - Go to the *Taxon lives in Habitat* tab
- Which microbial organisms can be found in a given food or habitat?
 - Go to the *Habitat is inhabited by Taxon* tab
- What are the phenotypes of a given microbe?
 - Go to the *Taxon exhibits Phenotype* tab
- Which microbes have this phenotype?
 - Go to the *Phenotype is exhibited by Taxon* tab
- In which use is studied this microbe?
 - Go to the *Taxon studied for Use* tab
- Which microbe is involved in this use?
 - Go to the *Use involves Taxon* tab

API

The screenshot shows the Swagger UI for the Florilège API. The interface includes a search bar with 'florilege-api.yaml' and an 'Explore' button. Below the search bar, the API is identified as 'Florilège API 0.0.1 OAS3'. A description states: 'API to access data in Florilège, a reference database of microorganism biodiversity.' A dropdown menu for 'Servers' is set to 'http://misaite.jouy.inra.fr/florilege-api/api - production server'. A 'default' dropdown is also visible. The main area lists several API endpoints:

Method	Endpoint	Description	Response
GET	/get/veersion	Returns version information about this instance.	getVersions
GET	/get/taxon/{taxid}	Returns the properties of a taxon given a NCBI taxon identifier.	getTaxon
GET	/get/obt/{obtId}	Returns the properties of an OntoBiotope concept given an OBT identifier.	getOBTConcept
GET	/get/doc/{sourcename}/{docid}	Returns a URL that describes a document.	getDocURL
GET	/search/taxon	Searches for taxa.	searchTaxon

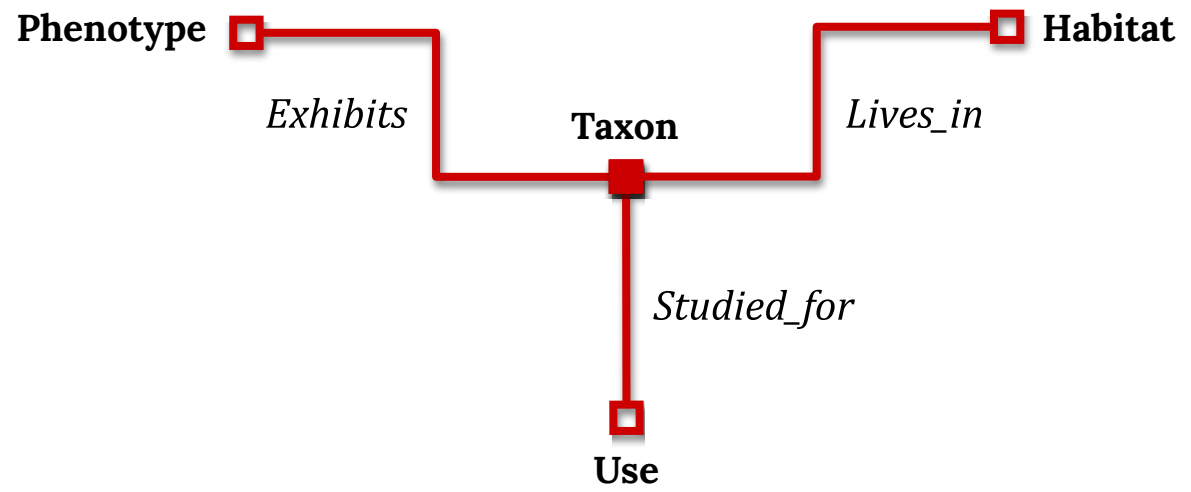


1

Le contenu de Florilège



Schéma entités-relations de Florilège



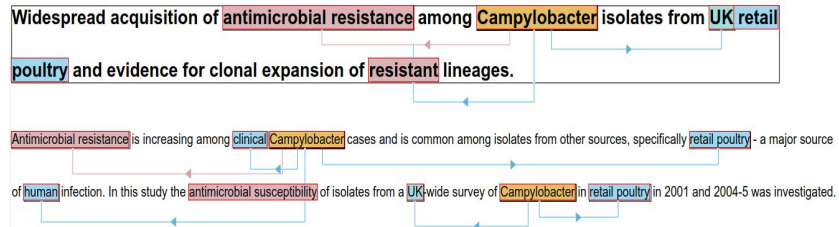


Florilège est une base intégrative

Catalogue DSMZ et CIRM

Bacillus coagulans	
DSMZ	
BACTERIA	How to read the following data (Example)
Name:	<i>Bacillus coagulans</i> Hammer 1915 emend. De Clerck et al. 2004
DSM No.:	1, Type strain
Strain designation:	609, CCM 2013, NCIB 9365, NCTC 10334, NRS 609, WDCM 00002
Other collection no. or WDCM no.:	CCM 2013, NCIB 9365, NCTC 10334, WDCM 00002
Isolated from:	evaporated milk
Country:	country of origin unknown
Date of sampling:	before 17.04.1974
Nagoya Protocol Restrictions:	There are NO known Nagoya Protocol restrictions for this strain.
History:	<- ATCC <- N.R. Smith, 609 <- J. Porter <- B. Hammer
Genbank accession numbers:	whole genome shotgun sequence: ALAS00000000 16S rRNA gene: DQ297928
Cultivation conditions:	Medium 453 , 40°C or Medium 1 , 40°C
Complete DSMZ Media List	

Résumés PubMed



Entrées GenBank

FEATURES

source

Location/Qualifiers

1..500

/organism="Bacillus coagulans"

/mol_type="genomic DNA"

/strain="lilac-01"

/isolation_source="the petals of Syringa vulgaris"

/db_xref="taxon:1398"

/country="Japan:Hokkaido, Asahikawa"

/collection_date="2010-06-03"

/collected_by="Kimiko Minamida"

<1..500

/product="16S ribosomal RNA"

rRNA



Florilège utilise des référentiels partagés

OntoBiotope

Ontologie des habitats, phénotypes
et usages des microbes.

```
[Term]
id: OBT:000050
name: artificial water environment
is_a: OBT:000006 ! artificial environment
```

```
[Term]
id: OBT:000051
name: asexual reproduction
synonym: "asexual growth" EXACT [TyDI:50000]
is_a: OBT:000018 ! phenotype wrt growth
```

```
[Term]
id: OBT:000052
name: atmosphere part
is_a: OBT:000013 ! natural environment habitat
```

```
[Term]
id: OBT:000053
name: autoploid
is_a: OBT:000022 ! phenotype wrt ploidy
```

<http://agroportal.lirimm.fr/ontologies/ONTOBIOTOPE>

- ▼ ↔ microbial habitat
 - ▶ ↔ animal husbandry and agricultural habitat
 - ▶ ↔ artificial environment
 - ▶ ↔ part of living organism
 - ▶ ↔ medical environment
 - ▶ ↔ microorganism associated habitat
 - ▶ ↔ planet
 - ▶ ↔ living organism
 - ▶ ↔ experimental medium
 - ▶ ↔ natural environment habitat
 - ▶ ↔ food
 - ▶ ↔ habitat wrt chemico-physical property
 - ▶ ↔ aquaculture habitat
 - ▶ ↔ animal habitat
- ▼ ↔ microbial phenotype
 - ▶ ↔ phenotype wrt genetic
 - ▶ ↔ phenotype wrt ploidy
 - ▶ ↔ phenotype wrt community behaviour
 - ▼ ↔ phenotype wrt stress
 - ▶ ↔ stress tolerant
 - ▶ ↔ phenotype wrt radiation impact
 - ▶ ↔ phenotype wrt relative humidity impact
 - ▶ ↔ **phenotype wrt chemical composition (2)**
 - ▶ ↔ phenotype wrt temperature impact
 - ▶ ↔ phenotype wrt molecule impact
 - ▶ ↔ stress sensitive
 - ▼ ↔ stress resistant
 - ↔ **metal resistant (2)**
 - ▶ ↔ **antibiotic resistant (2)**
 - ▶ ↔ **radiation resistant (2)**
 - ↔ **desiccation resistant (2)**
 - ↔ **alkali resistant (2)**
 - ↔ **acido resistant (2)**



Florilège utilise des référentiels partagés

NCBI Taxonomy

Taxonomie des taxons mentionnés dans les bases du NCBI.

2	prokaryote	prokaryote <Bacteria>	in-part	
2	prokaryotes	prokaryotes <Bacteria>	in-part	
6	Azorhizobium	scientific name		
6	Azorhizobium Dreyfus et al. 1988	emend. Lang et al. 2013		authority
7	ATCC 43989	ATCC 43989 <type strain>	type material	
7	Azorhizobium caulinodans	scientific name		
7	Azorhizobium caulinodans Dreyfus et al. 1988		authority	
7	Azotirhizobium caulinodans	equivalent name		
7	CCUG 26647	CCUG 26647 <type strain>	type material	
7	DSM 5975	DSM 5975 <type strain>	type material	
7	IFO 14845	IFO 14845 <type strain>	type material	
7	JCM 20966	JCM 20966 <type strain>	type material	
7	LMG 6465	LMG 6465 <type strain>	type material	
7	LMG:6465	LMG:6465 <type strain>	type material	
7	NBRC 14845	NBRC 14845 <type strain>	type material	
7	ORS 571	ORS 571 <type strain>	type material	
9	Acyrtosiphon pisum symbiont P		includes	
9	Buchnera aphidicola	scientific name		
9	Buchnera aphidicola Munson et al. 1991		authority	
9	primary endosymbiont of Schizaphis graminum		includes	
10	"Cellvibrio" Winogradsky 1929		authority	
10	Cellvibrio	scientific name		
10	Cellvibrio (ex Winogradsky 1929)	Blackall et al. 1986	emend. Humphry et al. 2003	

Lineage (full): [cellular organisms](#); [Bacteria](#); [Terrabacteria group](#)

- [Firmicutes](#) *Click on organism name to get more information.*
 - [Bacilli](#)
 - [Clostridia](#)
 - [Culicoidibacteria](#)
 - [Erysipelotrichia](#)
 - [Limnochordia](#)
 - [Negativicutes](#)
 - [Thermolithobacteria](#)
 - [Tissierella](#)
 - [Firmicutes sensu stricto incertae sedis](#)
 - [unclassified Firmicutes sensu stricto](#)
 - [environmental samples](#)

<https://www.ncbi.nlm.nih.gov/taxonomy>



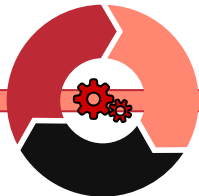
Florilège est alimentée par des techniques de text-mining



Sources



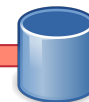
Documents



Text-mining



Documents
annotés



Base de
données



Services

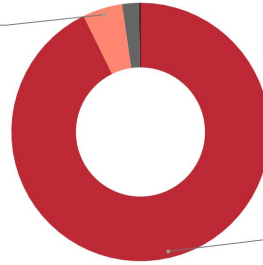


Florilège est gros

Documents

29 120 000

GenBank
5.0%

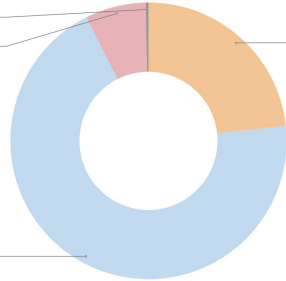


PubMed
92.7%

Entités
40 836 000

Use
0.3%

Phenotype
7.1%



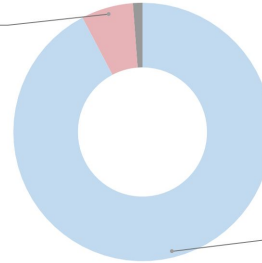
Taxon
23.3%

Habitat
69.3%

Relations

9 751 000

Exhibits
6.5%



Lives_in
92.3%



2

L'API de Florilege



Motivations

- ◆ Mettre les données à disposition pour les bioinformaticiens.
- ◆ Permettre des usages nouveaux:
 - ◆ Lier avec d'autres bases de données;
 - ◆ Construire des visualisations.



Motivations

Path source

- Bacteria
- Terrabacteria group
- Actinobacteria
- Actinobacteria

- Candidatus Nanopelagiales
- Glycomycetales
- Geodermatophilales
- Kineosporiales
- Frankiales
- Micromonosporales
- Streptomycetales
- Streptosporangiales
- Pseudonocardiales
- Acidothermales
- Jiangellales
- Actinopolysporales
- Propionibacteriales
- unclassified Actinobacteria (class)
- Micrococcales
- Corynebacteriales
- Actinomycetales
- Bifidobacteriales

- edible film
- plant product and primary derivative thereof
- additive
- composite food
- animal product and primary derivative thereof
- liquid food
- food for particular diet

Path target

- microbial habitat
- food
- food for human
- commodity and primary derivative thereof

<http://bibliome.opscidia.com/>



Documentation de l'API (OpenAPI)

Florilege API ^{0.0.1} OAS3

florilege-api.yaml

API to access data in Florilege, a reference database of microorganism biodiversity.

Servers

<http://migale.jouy.inra.fr/florilege-api/api> - production server

default

GET	/get/version	Returns version information about this instance.	getVersions
GET	/get/taxon/{taxid}	Returns the properties of a taxon given a NCBI taxon Identifier.	getTaxon
GET	/get/obt/{obtid}	Returns the properties of an OntoBiotope concept given an OBT Identifier.	getOBTConcept
GET	/get/doc/{sourcename}/{docid}	Returns an URL that describes a document.	getDocURL
GET	/search/taxon	Searches for taxa.	searchTaxon
GET	/search/obt	Searches for OntoBiotope concepts.	searchOBTConcept
GET	/search/relations	Searches for relations.	searchRelations
GET	/search/join-relations	Searches for joined relations.	searchJoinRelations

<http://migale.jouy.inra.fr/florilege-api/api-doc/>



Fonctions recherche d'entités

Taxon

Habitat Phenotype Use

`/get/taxon/{taxid}`

`/get/obt/{obtid}`

Get

```
{
  "id": "ncbi:1423",
  "name": "Bacillus subtilis",
  "Path": ["ncbi:1/ncbi:131567/ncbi:2/.../ncbi:1423"],
  "qps": "yes"
}
```

```
{
  "id": "OBT:000427",
  "type": "habitat",
  "name": "soil",
  "synonyms": ["soilbourne"],
  "path": ["OBT:000001/OBT:000013/OBT:000158/OBT:000427"]
}
```

`/search/taxon`

`/search/obt`

Params: requête, ancêtre, QPS

Params: requête, ancêtre, type

Search

```
[
  {"id": "ncbi:1423"},
  {"id": "ncbi:653685"},
  {"id": "ncbi:1386"},
  ...
]
```

```
[
  {"id": "OBT:000427"},
  {"id": "OBT:000158"},
  {"id": "OBT:000013"},
  ...
]
```




Fonction recherche de relations

`/search/relations`

Paramètres:

- Identifiant du taxon
- Identifiant du habitat/phenotype/use
- Type de l'entité
- Statut QPS du taxon
- Sources (PubMed, GenBank, DSMZ, CIRM)

```
[  
  {  
    "taxid": "ncbi:909932",  
    "obtid": "OBT:000397",  
    "type": "habitat",  
    "source": "PubMed",  
    "docs": ["11589654"],  
    "taxon_forms": ["Negativicutes"],  
    "obt_forms": ["prepared food"]  
  },  
  {  
    "taxid": "ncbi:1737404",  
    "obtid": "OBT:000339",  
    "type": "habitat",  
    "source": "PubMed",  
    "docs": ["23687466"],  
    "taxon_forms": ["Tissierellia"],  
    "obt_forms": ["meal"]  
  },  
  {  
    "taxid": "ncbi:186801",  
    "obtid": "OBT:000477",  
    "type": "habitat",  
    "source": "PubMed",  
    "docs": ["14589747"],  
    "taxon_forms": ["Clostridia"],  
    "obt_forms": ["additive"]  
  },  
  ...  
]
```



Fonction recherche de jointures

**Quels microorganismes
partagent les mêmes
habitats ?**



**Quels habitats présentent
des flores communes ?**





Fonction recherche de jointures

`/search/relations`

Paramètres :

- Types des entités à gauche, à droite et en jointure
- Identifiants des entités à gauche, à droite et en jointure
- Statut QPS du taxon
- Sources (PubMed, GenBank, DSMZ, CIRM)

```
[
  {
    "leftType" : "habitat",
    "leftRoot" : "OBT:000003",
    "leftId" : "OBT:001620",
    "leftSource" : "PubMed",
    "leftDocs" : ["21313850"],
    "rightType" : "habitat",
    "rightRoot" : "OBT:000004",
    "rightId" : "OBT:000008",
    "rightSource" : "PubMed",
    "rightDocs" : ["20113566", "2554624", "24047612"],
    "joinType" : "taxon",
    "joinId" : "ncbi:114727"
  },
  {
    "leftType" : "habitat",
    "leftRoot" : "OBT:000003",
    "leftId" : "OBT:001620",
    "leftSource" : "PubMed",
    "leftDocs" : ["21313850"],
    "rightType" : "habitat",
    "rightRoot" : "OBT:000004",
    "rightId" : "OBT:003316",
    "rightSource" : "PubMed",
    "rightDocs" : ["18414973", "25520513", "26629623"],
    "joinType" : "taxon",
    "joinId" : "ncbi:114727"
  },
  ...
]
```



Utiliser l'API

- ◆ Contactez-nous
 - ◆ florilege@inrae.fr
- ◆ RTFM
 - ◆ <http://migale.jouy.inra.fr/florilege-api/api-doc/>



Utiliser l'API

- ◆ Contactez-nous
 - ◆ florilege@inrae.fr
- ◆ RTFM
 - ◆ <http://migale.jouy.inra.fr/florilege-api/api-doc/>



Utiliser l'API

- ◆ Contactez-nous
 - ◆ florilege@inrae.fr
- ◆ RTFM
 - ◆ <http://migale.jouy.inra.fr/florilege-api/api-doc/>

MERCI



Indicateurs de qualité

- ◆ Certaines sources sont plus fiables
- ◆ Mesures de confiance des outils de text-mining
- ◆ Projet: capture de feedback



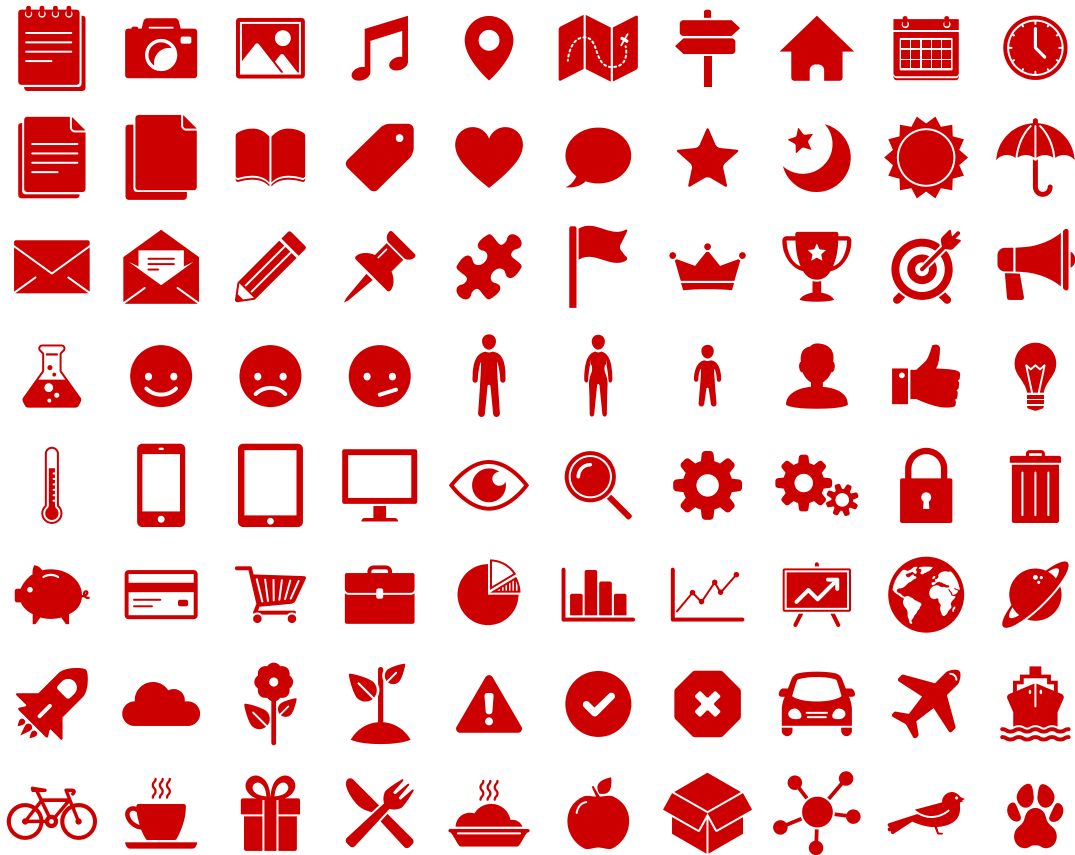
Mise à jour

- ◇ *ad hoc*, beaucoup de dépendances
 - ◇ Sources
 - ◇ Référentiels
 - ◇ Technologies d'extraction
- ◇ *Projet*: māj régulières



Nouvelles sources, nouveaux référentiels

- ◆ OntoBiotope en évolution continue.
- ◆ NCBI Taxonomy à compléter (niveaux subspécifiques et souches).
- ◆ En projet: GOLD, EPMC.



SlidesCarnival icons are editable shapes.

This means that you can:

- Resize them without losing quality.
- Change fill color and opacity.
- Change line color, width and style.

Isn't that nice? :)

Examples:

