

GFBio Terminology Service

Enabling a Research Data Management beyond Data Heterogeneity

Prof. Dr. Claudia Müller-Birn, Dr. Naouel Karam, Robert Lorenz

Institute for Computer Science, Freie Universität Berlin

E-Science-Tage 2017: Forschungsdaten managen

16.03.2017

Data Repositories in Biology and Biodiversity Research



Data Repositories in Biology and Biodiversity Research



Different Disciplines/Communities

Different Data Formats

Different Access Rights



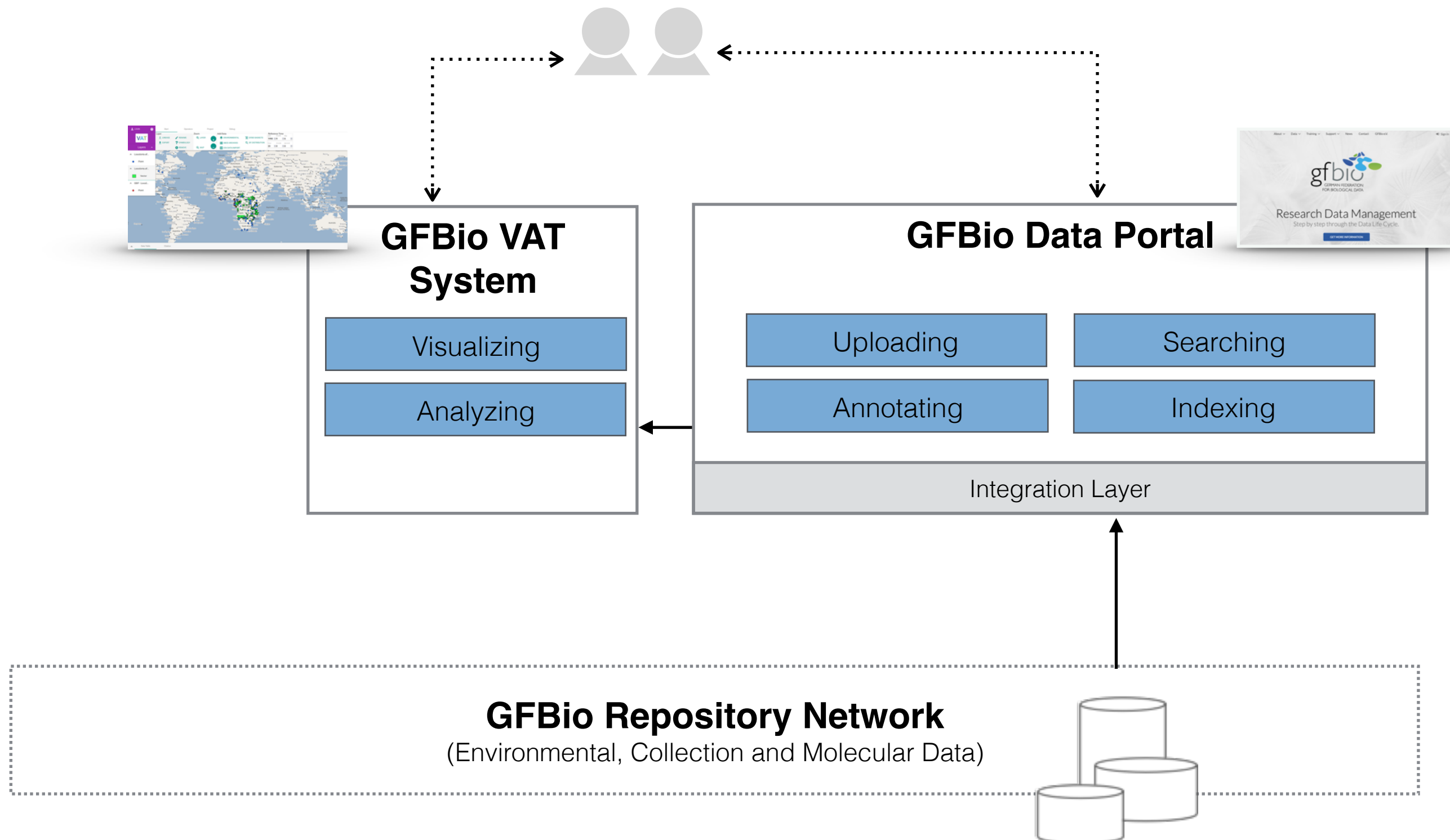
museum für
naturkunde
berlin



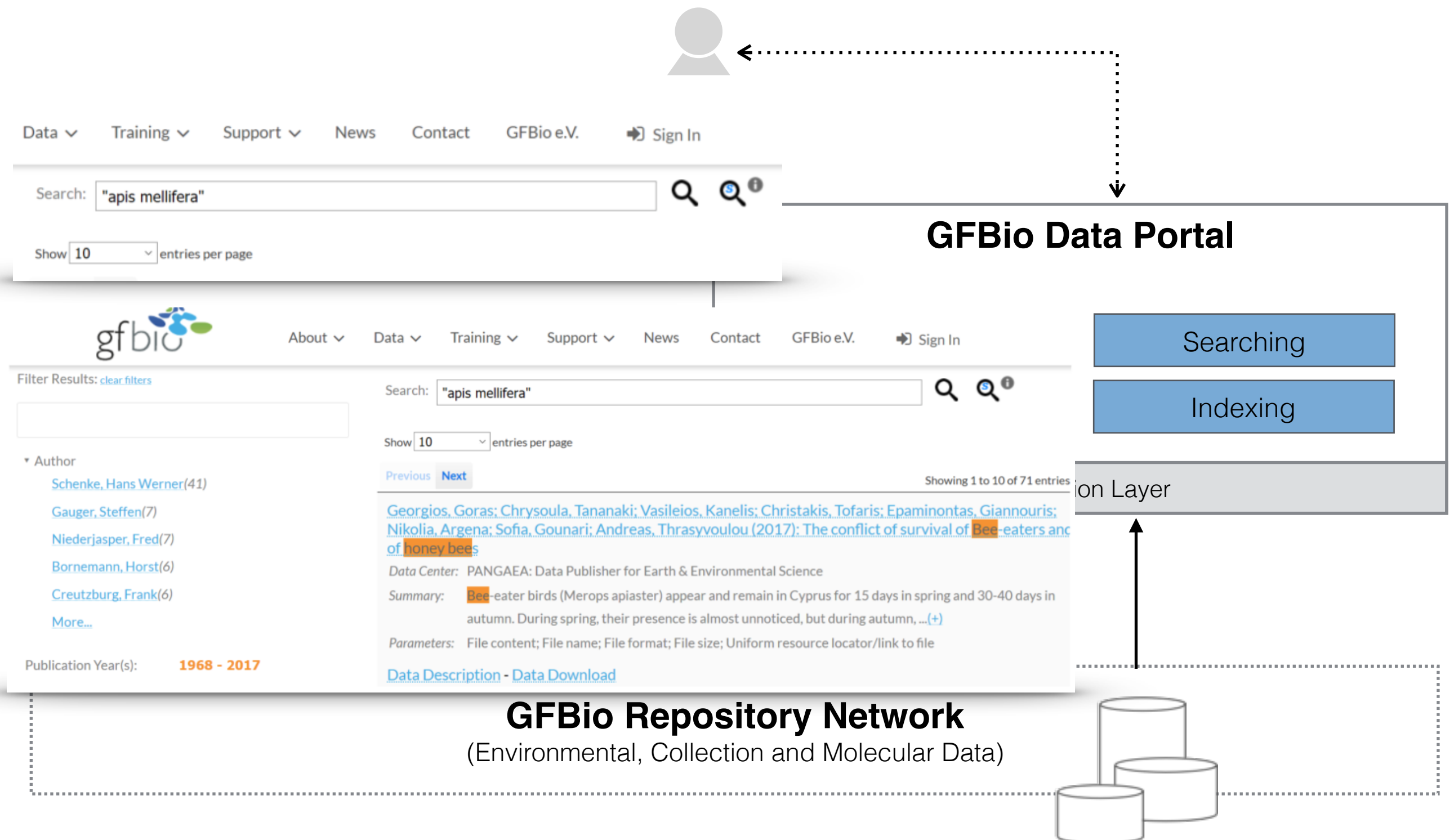
NATURKUNDE
MUSEUM
STUTTGART



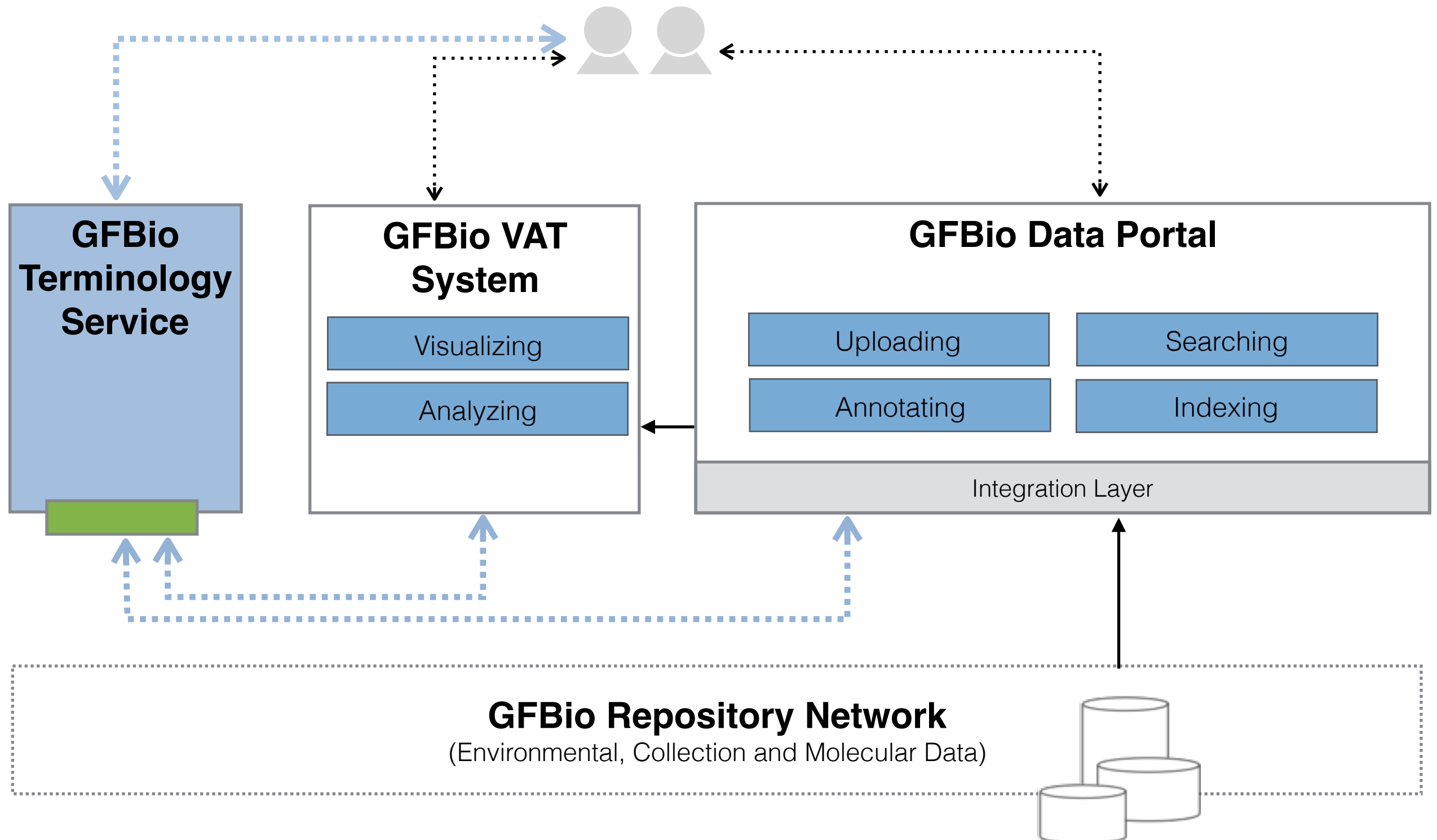
The GFBio Infrastructure



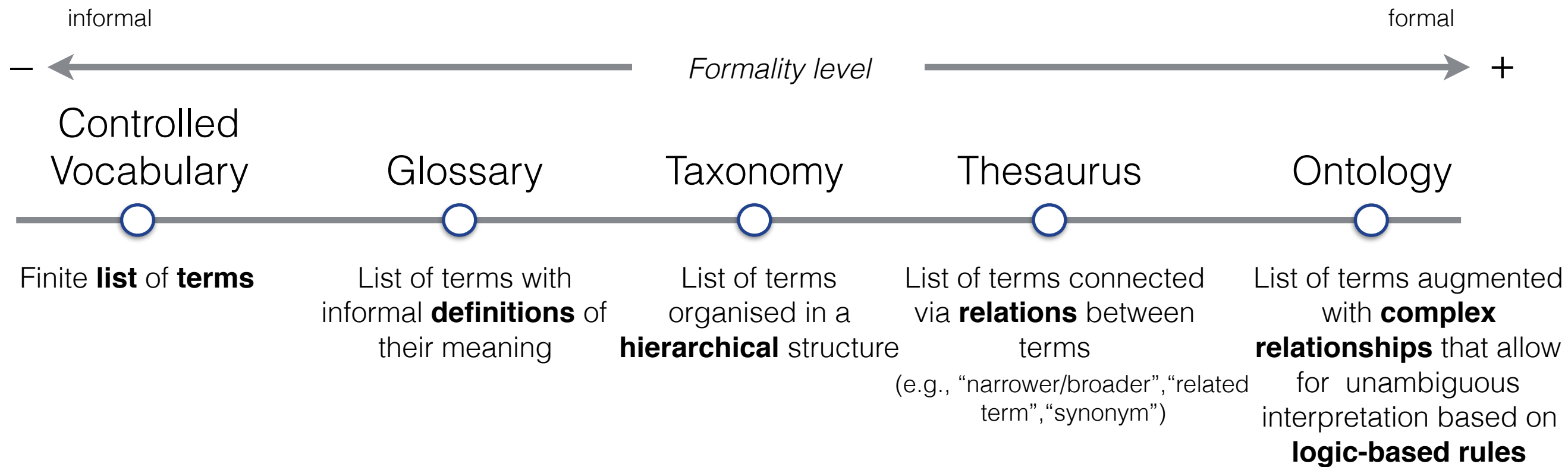
The GFBio Infrastructure



The Terminology Service within GFBio Infrastructure



Terminology Formality Levels



water

definition [type: xsd:string]

An oxygen hydride consisting of an oxygen atom that is covalently bonded to two hydrogen atoms.

chemical entity

molecular entity

inorganic molecular entity

inorganic hydride

chalcogen hydride

oxygen hydride

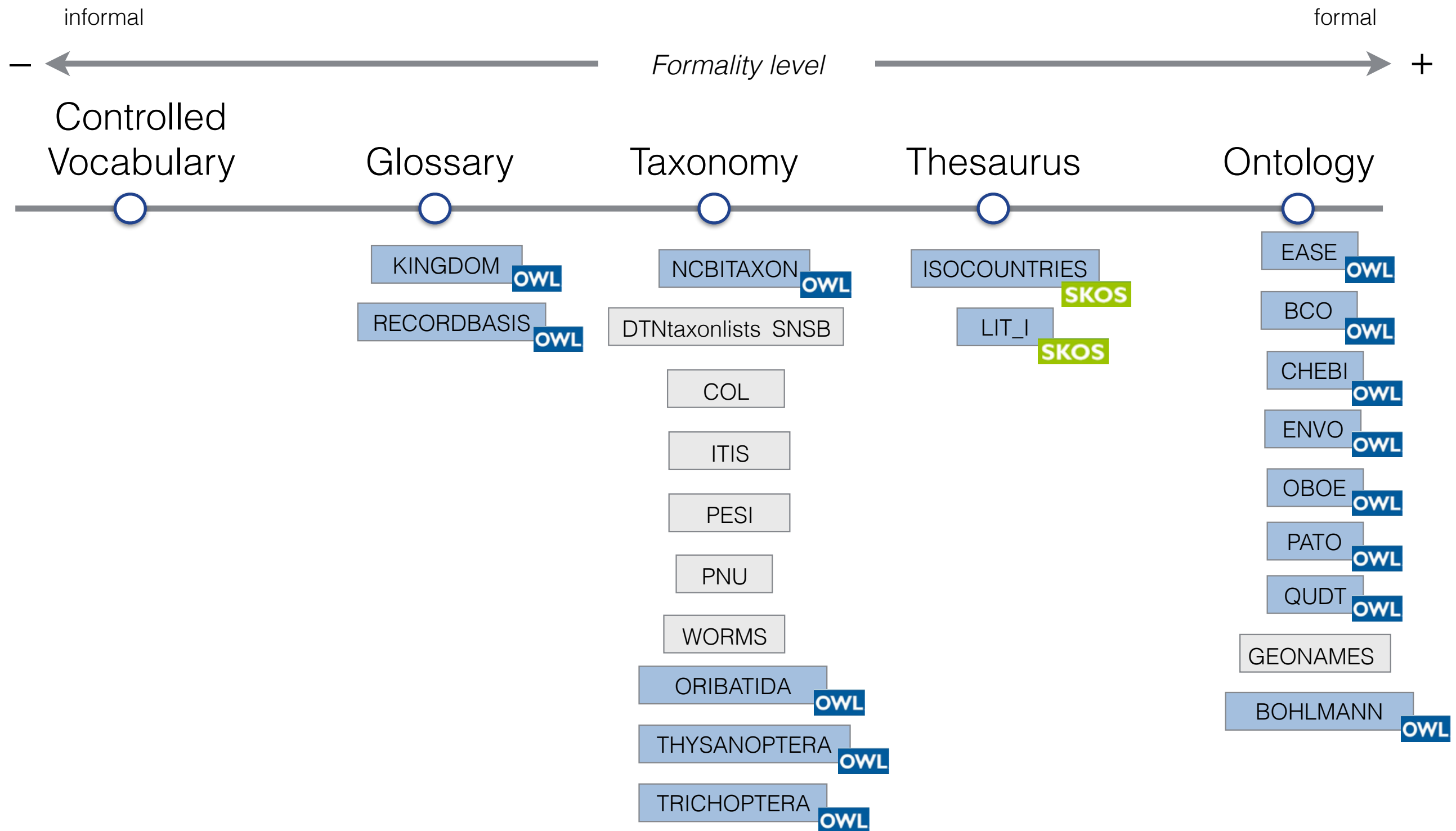
water

has_exact_synonym [type: xsd:string]

oxidane

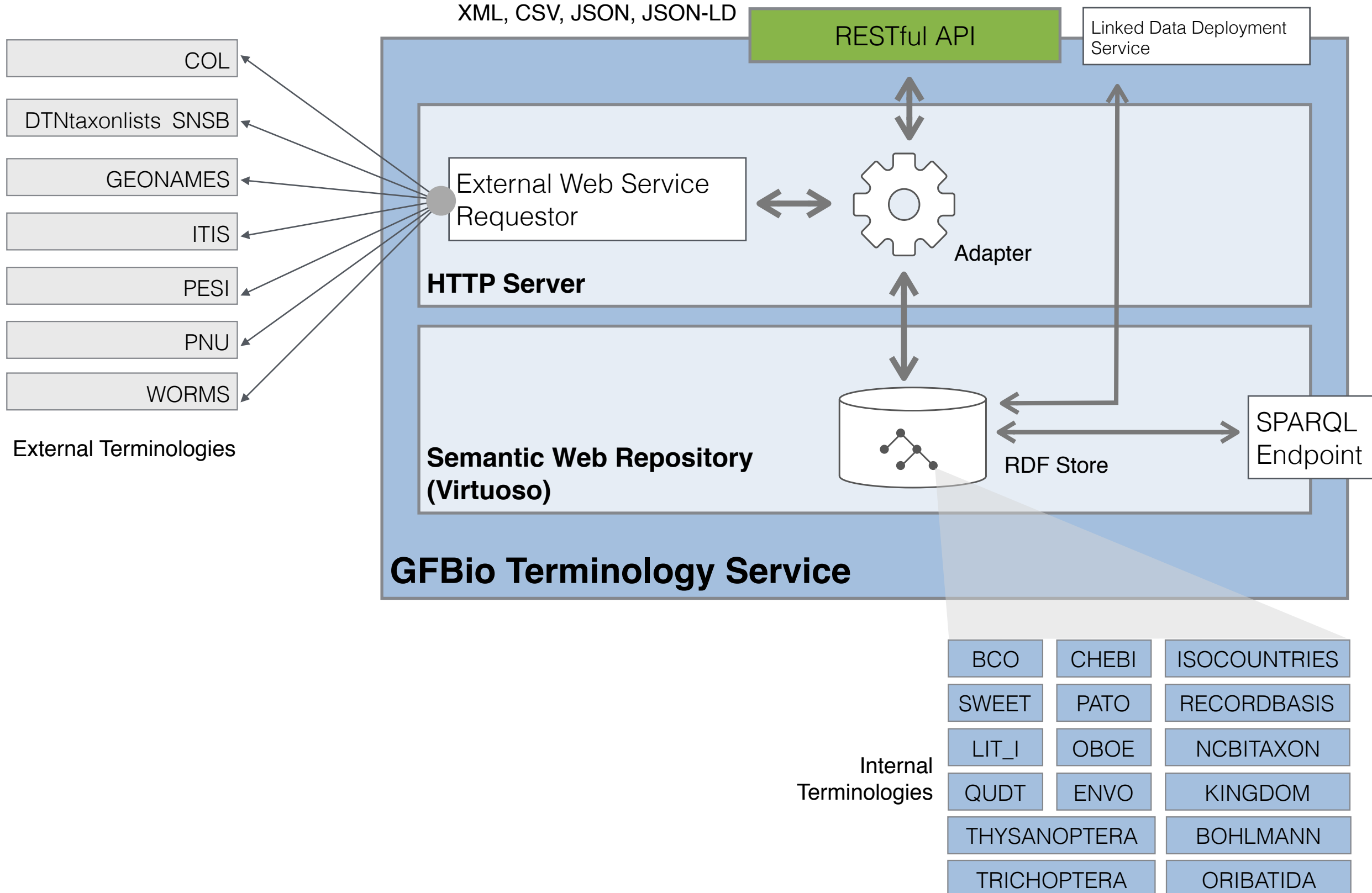
- **has role** some greenhouse gas
- **has role** some mouse metabolite
- **is conjugate acid of** some hydroxide
- **has role** some amphiprotic solvent
- **has role** some *Saccharomyces cerevisiae* metabolite
- **has role** some *Escherichia coli* metabolite
- **has role** some human metabolite
- **is conjugate base of** some oxonium

Examples from Terminology Service



The Terminology Service Architecture

Output formats:
XML, CSV, JSON, JSON-LD



The Terminology Service API

BROWSE

ACCESS

CONSUME

CONTRIBUTE

Search endpoints

GET /terminologies/**search**
GET terminologies/**suggest**

Terminology endpoints

GET /terminologies
GET /terminologies/<terminology_id>
GET /terminologies/<terminology_id>/**metrics**
GET /terminologies/<terminology_id>/**metadata**

Term endpoints

GET /terminologies/<terminology_id>/**allterms**
GET /terminologies/<terminology_id>/**term**
GET /terminologies/<terminology_id>/**synonyms**

Hierarchy-specific endpoints

GET /terminologies/<terminology_id>/**narrower**
GET /terminologies/<terminology_id>/**allnarrower**
GET /terminologies/<terminology_id>/**broader**
GET /terminologies/<terminology_id>/**allbroader**
GET /terminologies/<terminology_id>/**hierarchy**

<http://terminologies.gfbio.org/api/>

RESTful API

GFBio Terminology Service

Researchers browse Terminologies



BROWSE terminologies that are relevant to your research in order to understand how to use them.

carbon dioxide CHEBI

http://purl.obolibrary.org/obo/CHEBI_16526

Description

A one-carbon compound with formula CO₂ in which the carbon is attached to each oxygen atom by a double bond. A colourless, odourless gas under normal conditions, it is produced during respiration by all animals, fungi and microorganisms that depend directly or indirectly on living or decaying plants for food.

Synonyms

CO₂, CARBON DIOXIDE, Carbon dioxide, E 290, E-290, E290, R-744, [CO₂], carbon dioxide, carbonic anhydride, dioxidocarbon

Further Term Details

Term Relations

Sub Class Of

http://purl.obolibrary.org/obo/CHEBI_23014

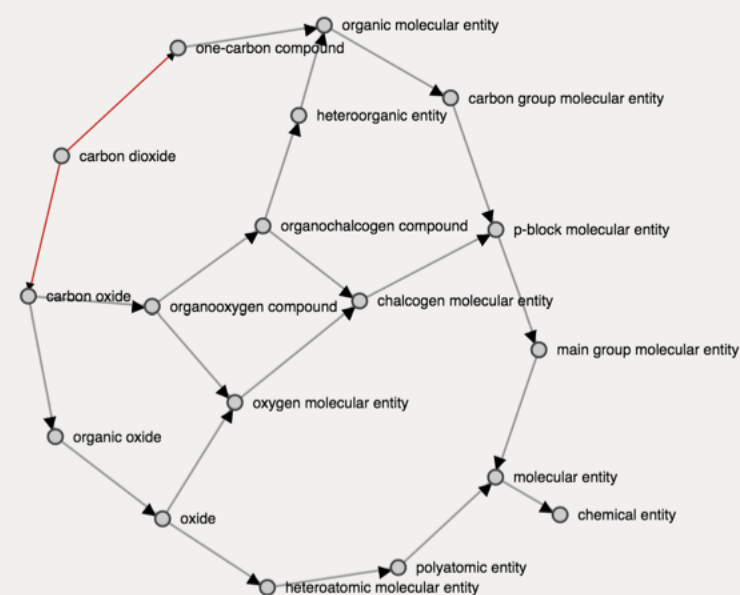
http://purl.obolibrary.org/obo/CHEBI_64708

Tree View

Network View



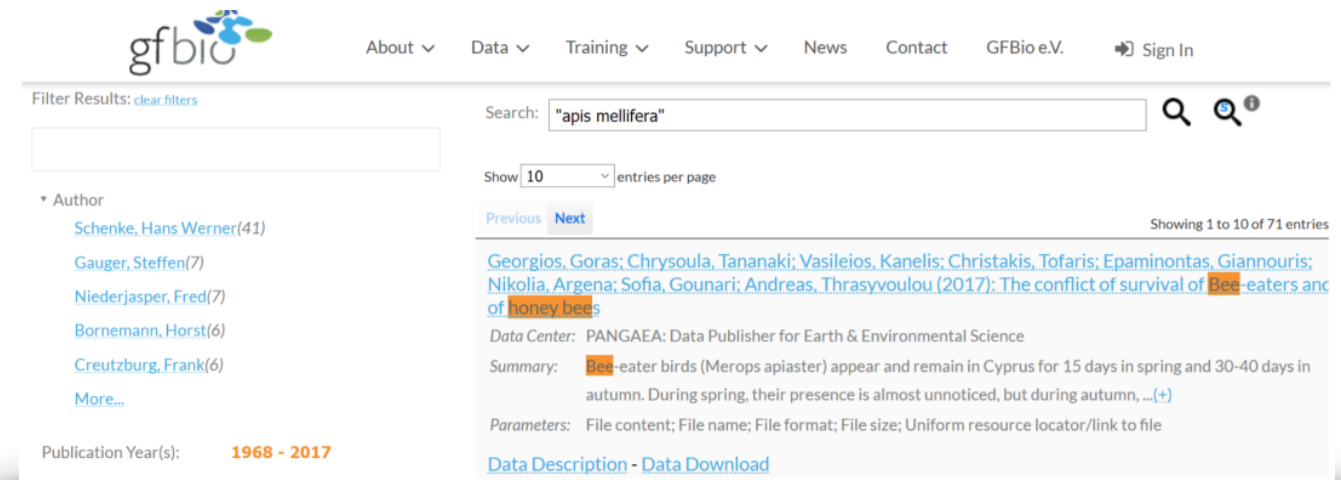
isOntology=true



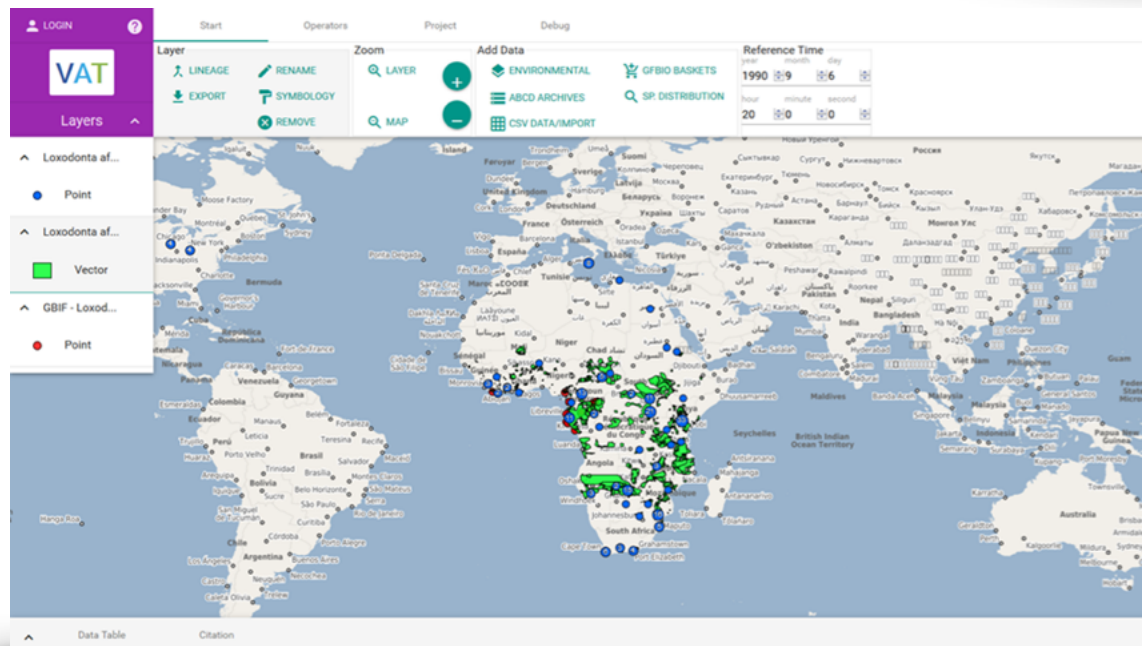
Developers build Semantic Services



ACCESS terminologies programmatically to provide semantically enriched web services.



The screenshot shows the gfbio web application. The top navigation bar includes links for About, Data, Training, Support, News, Contact, GFBio e.V., and Sign In. A search bar contains the text "apis mellifera". Below the search bar, there are filters for Author and Publication Year(s). The Author filter lists: Schenke, Hans Werner(41), Gauger, Steffen(7), Niederjasper, Fred(7), Bornemann, Horst(6), Creutzburg, Frank(6), and More... The Publication Year(s) filter shows 1968 - 2017. The search results section shows 10 entries per page. The first entry is by Georgios, Goras; Chrysoula, Tananaki; Vasileios, Kanelis; Christakis, Tofaris; Epaminontas, Giannouris; Nikolia, Argenta; Sofia, Gounari; Andreas, Thrasylvoulou (2017): The conflict of survival of Bee-eaters and of honey bees. The Data Center is PANGAEA: Data Publisher for Earth & Environmental Science. The Summary is: Bee-eater birds (Merops apiaster) appear and remain in Cyprus for 15 days in spring and 30-40 days in autumn. During spring, their presence is almost unnoticed, but during autumn, ...(+). The Parameters are: File content; File name; File format; File size; Uniform resource locator/link to file. There are links for Data Description and Data Download.



Developers download Terminologies



CONSUME information from the terminologies to your information system to use them locally.

```
{
  took: 1,
  timed_out: false,
  _shards: {
    total: 1,
    successful: 1,
    failed: 0
  },
  hits: {
    total: 1,
    max_score: 12.257323,
    hits: [
      {
        _index: "features_v1",
        _type: "feature",
        _id: "1047579",
        _score: 12.257323,
        _source: {
          internal-source: "standard",
          internal-datestamp: "2015-05-05T15:06:18.680Z",
          name: "Globigerina bulloides",
          ontology: 1,
          mainTopics: "Biological Classification",
          topics: [
            "Foraminifera",
            "Chromista",
            "Biological Classification"
          ],
          terms: [
            "Globigerina bulloides",
            "Globigerina",
            "Globigerininae",
            "Globigerinidae",
            "Globigerinoidea",
            "Globigerinina",
            "Rotaliida",
            "Globobulimina",
            "Foraminifera",
            "Rhizaria",
            "Haptophyta",
            "Chromista"
          ]
        }
      }
    ]
  }
}
```

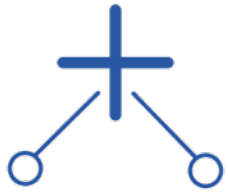


PANGAEA.

Data Publisher for Earth & Environmental Science

Extended metadata information
by
using broader terms
and synonyms
from the Terminology Service

Researchers share Terminologies



CONTRIBUTE your terminologies to the TS to share them with your research team and the world.



Uploading/Accessing terminologies



- DTNtaxonlists SNSB
- PNU
- EASE
- ISOCOUNTRIES
- KINGDOM
- LIT_I
- RECORDBASIS
- BOHLMANN
- ORIBATIDA
- THYSANOPTERA
- TRICHOPTERA

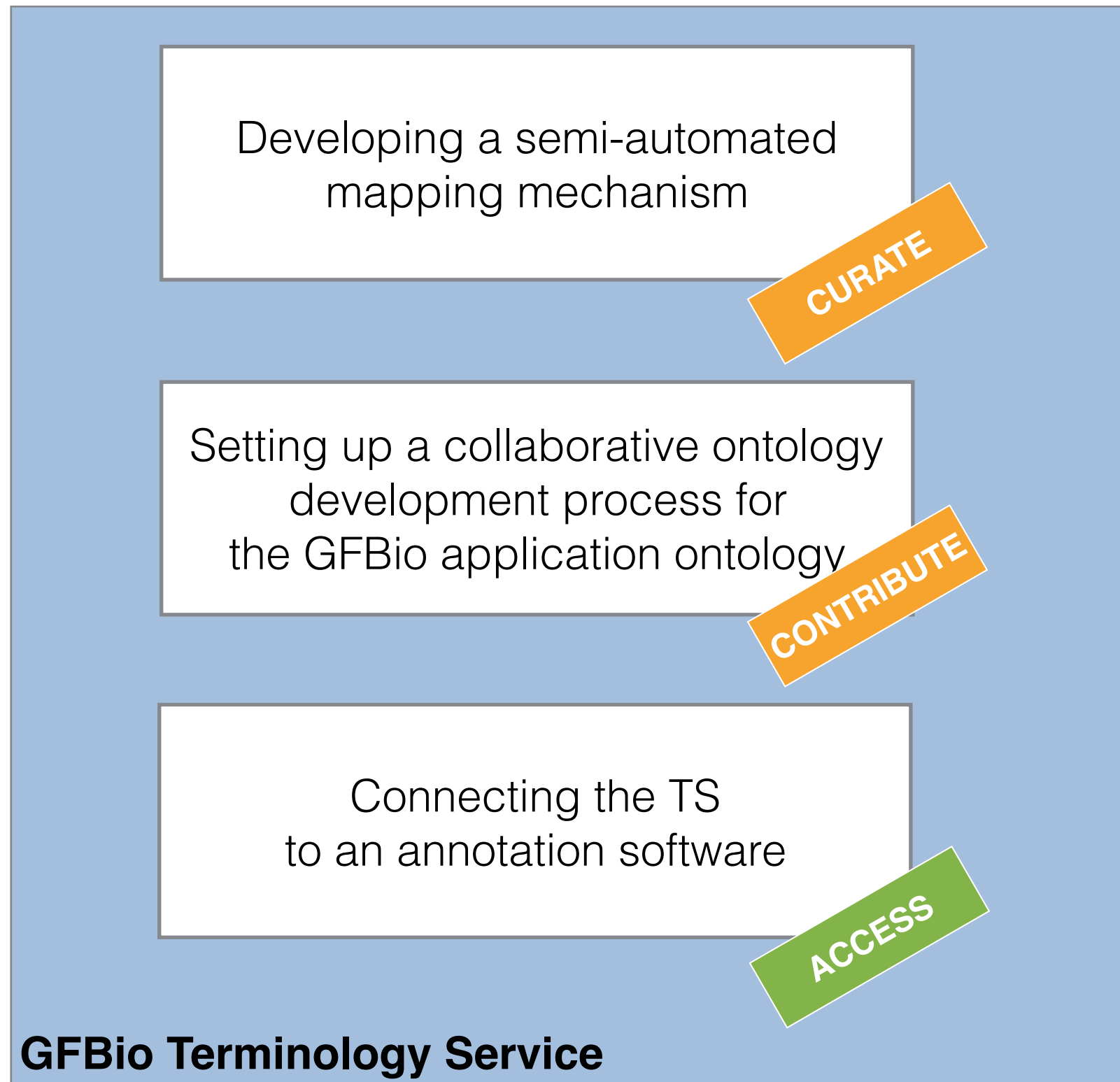
Suggesting terminologies

GFBio Community



Adding metadata of terminologies

Current Activities and Next Steps



GFBio Team @ E-Science Tage

Freitag 17.03.2017



11:30 - 14:00 Uhr Gesprächstisch (Foyer, EG)

GFBio e.V.

Tina Astor (SUB/Uni Göttingen), Janine Felden (Uni Bremen/MARUM), Ivo Kostadinov (Jaobes University)



16:30 - 17:30 Uhr Keynote (Aula, 2. OG)

GFBio - German Federation for Biological Data
Prof. Dr. GLÖCKNER, Frank Oliver

GFBio Terminology Service

Enabling a Research Data Management beyond Data Heterogeneity

Karam,N.; Müller-Birn,C.,Gleisberg,M,; Fichtmüller,D.; Tolksdorf,R.Güntsch, A. (2016): A Terminology Service supporting semantic annotation, integration, discovery and analysis of interdisciplinary research data. Datenbank-Spektrum. Zeitschrift für Datenbanktechnologien und Information Retrieval, Vol. 16, No.3, 195-205.