



i4Life



Freie Universität  Berlin

Using i4Life services in workflows

Anton Güntsch

Freie Universität Berlin

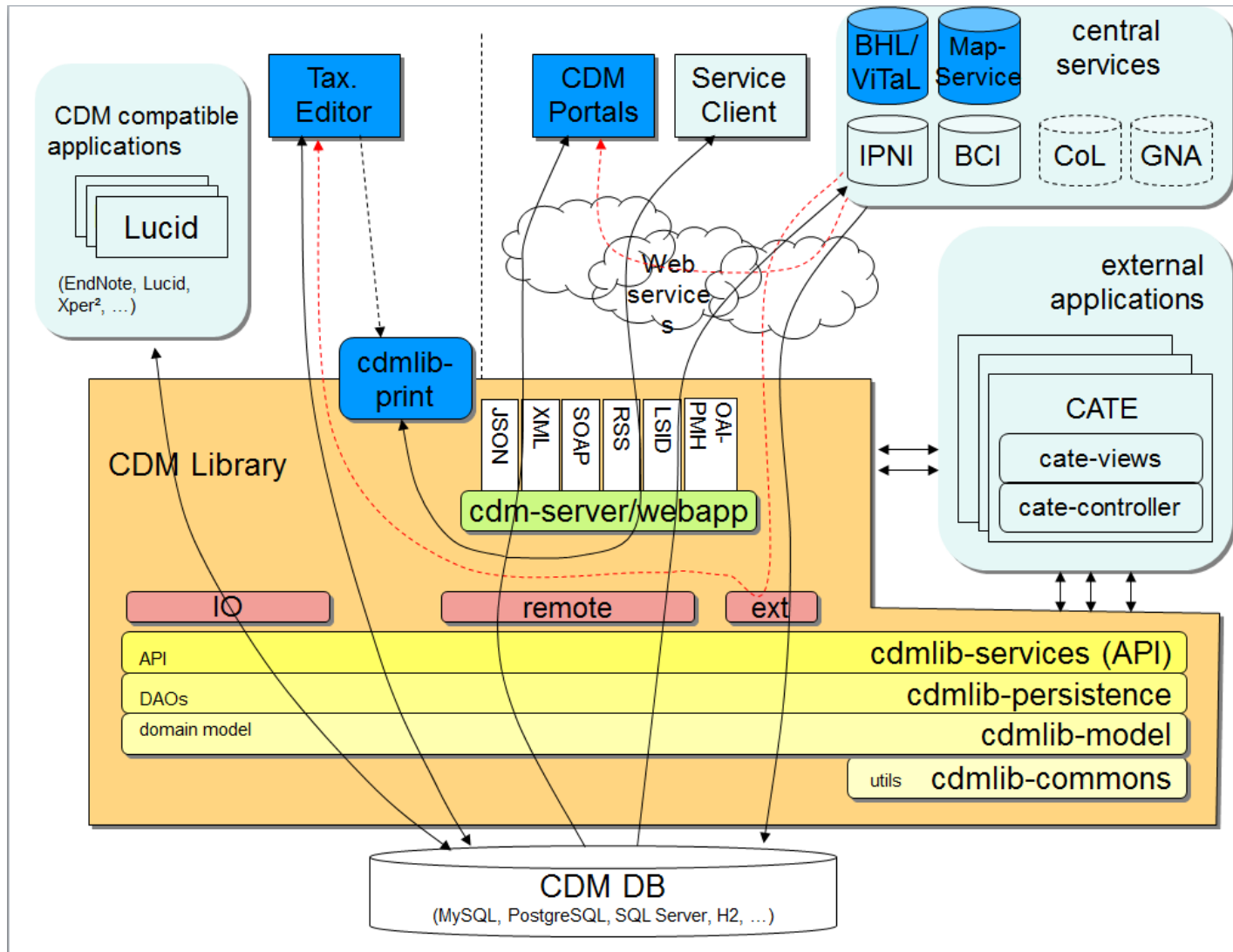
Botanic Garden & Botanical Museum Berlin-Dahlem

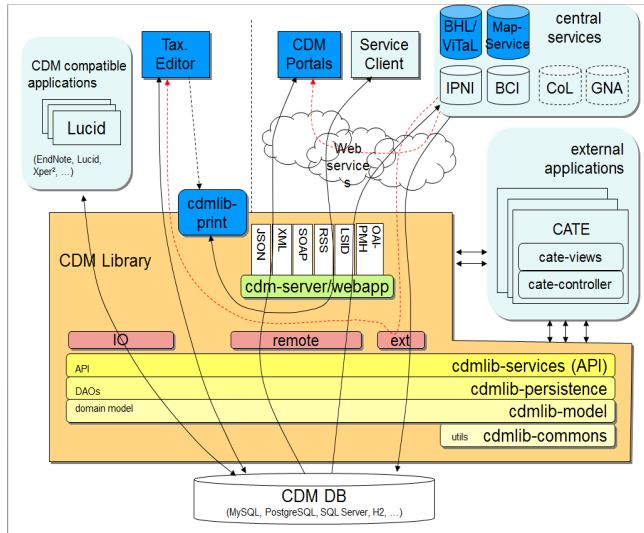
Dept. of Research and Biodiversity Informatics

BiodiversityInformatics@bgbm.org



The EDIT Platform for Cybertaxonomy as a global partner in i4Life

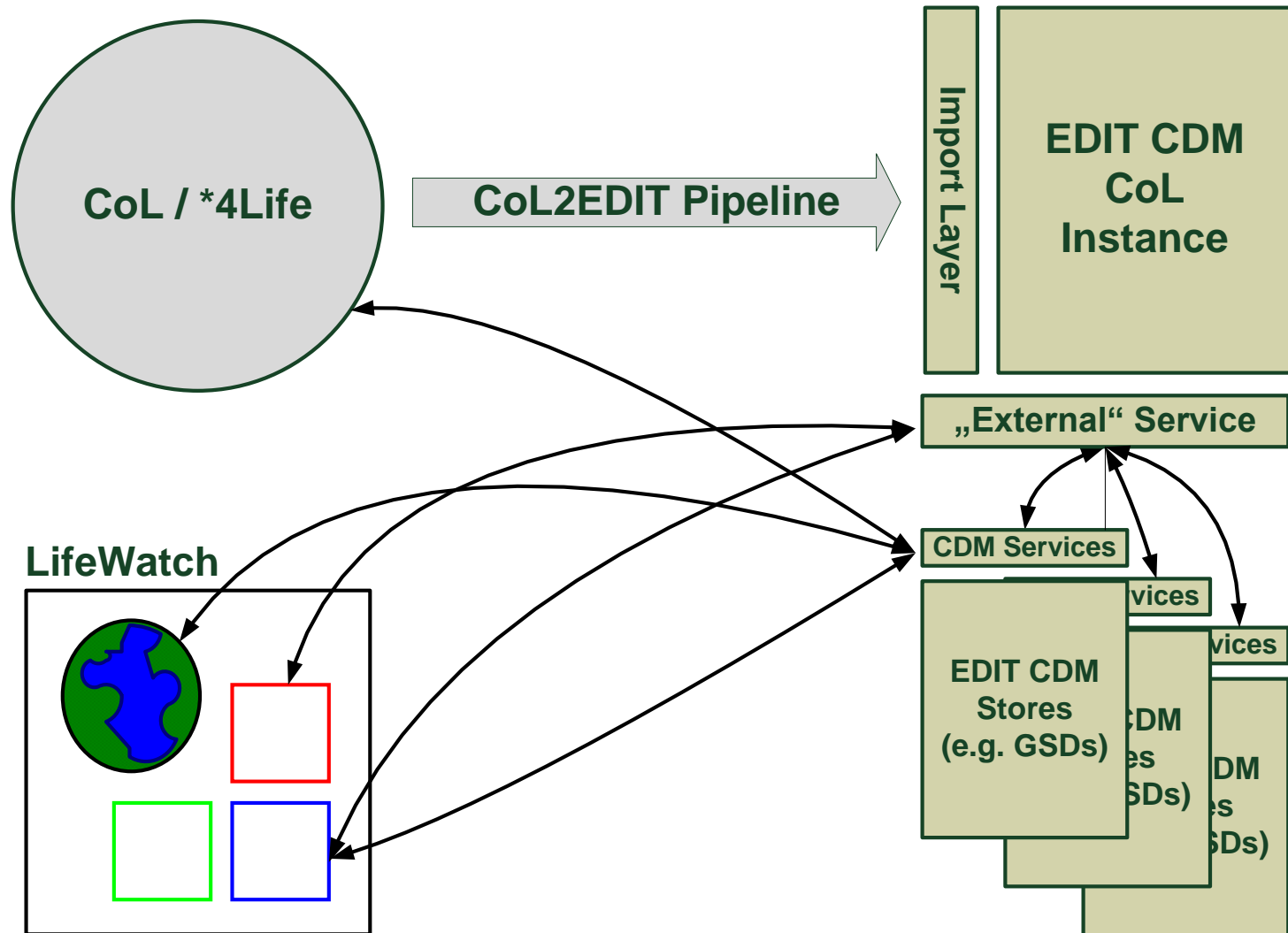




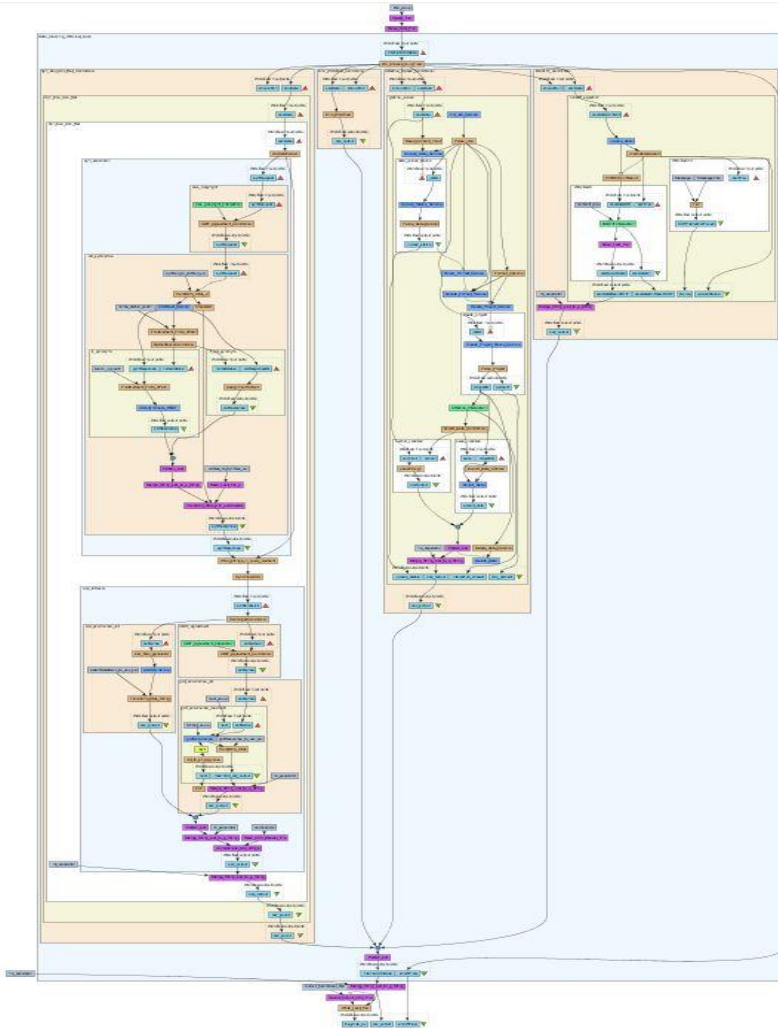
- The EDIT Platform for Cybertaxonomy is a comprehensive software environment supporting the taxonomic workflow from field work to printed publications.
- Based on the „Common Data Model“ (CDM)
- A growing number of software tools (portals, editors, import/export, pdf-generators, ...).

→ The EDIT Platform is not an infrastructure. It is used to build infrastructures.

The EDIT Platform for Cybertaxonomy as a global partner in i4Life



Target use case: BioVeL workflows



- EU 7th Framework project developing a “Biodiversity Virtual e-Laboratory”
- Uses workflows for the performance of biodiversity analysis tasks such as
 - Analysis of marine biodiversity changes based on the comparison of contemporary research data with museum records.
 - Ecological niche modeling of invasives in the Baltic.
- CoL Services integrated into „Data Refinement Workflow” used for preparing and cleaning source data sets.

Target use case: BioVeL workflows



Platform for Cybertaxonomy

[CDM Library](#) | [Taxonomic Editor](#) | [CDM Server](#) | [Dataportal](#) | [Print Publisher](#) | [CDM Setups](#)

CDM Library

General Information

- [What is the CDM Library?](#)
- [Details](#)
- [Releases](#)

Developer

- [Getting started](#)
- [Reference](#)
- [UML](#)
- [API](#)
- ▼ [REST API](#)
 - [Search](#)
 - [Statistics](#)
 - [Name Catalogue](#)

- [OAI-PMH](#)
- [Troubleshooting](#)
- [Wiki](#)

Project

- [Mailing Lists](#)
- [Project Team](#)
- [Continuous Integration](#)
- [Issue Tracking](#)
- [Source Repository](#)
- [License](#)

Name Catalogue REST API

This web service namespace is an add-on to the already existing CDM REST API and provides information relating to scientific names as well as taxa present in the underlying datasource.

The services include,

- [name search](#) : This web service endpoint serves as a search engine for scientific name-related taxonomic information.
- [name information](#) : This web service endpoint provides information related a specific name uuid.
- [taxon information](#) : This web service endpoint provides information related a specific taxon uuid.
- [accepted name search](#) : This web service endpoint serves as a search engine for retrieving taxonomic information of the accepted name of the synonymy to which the queried scientific name belongs.
- [fuzzy name search](#) : This web service endpoint serves as a fuzzy search for taxonomic names.

These web services have been developed within the i4Life project (<http://www.i4life.eu>) using data imported from the Catalogue of Life (<http://www.catalogueoflife.org>).

CDM - Catalogue of Life

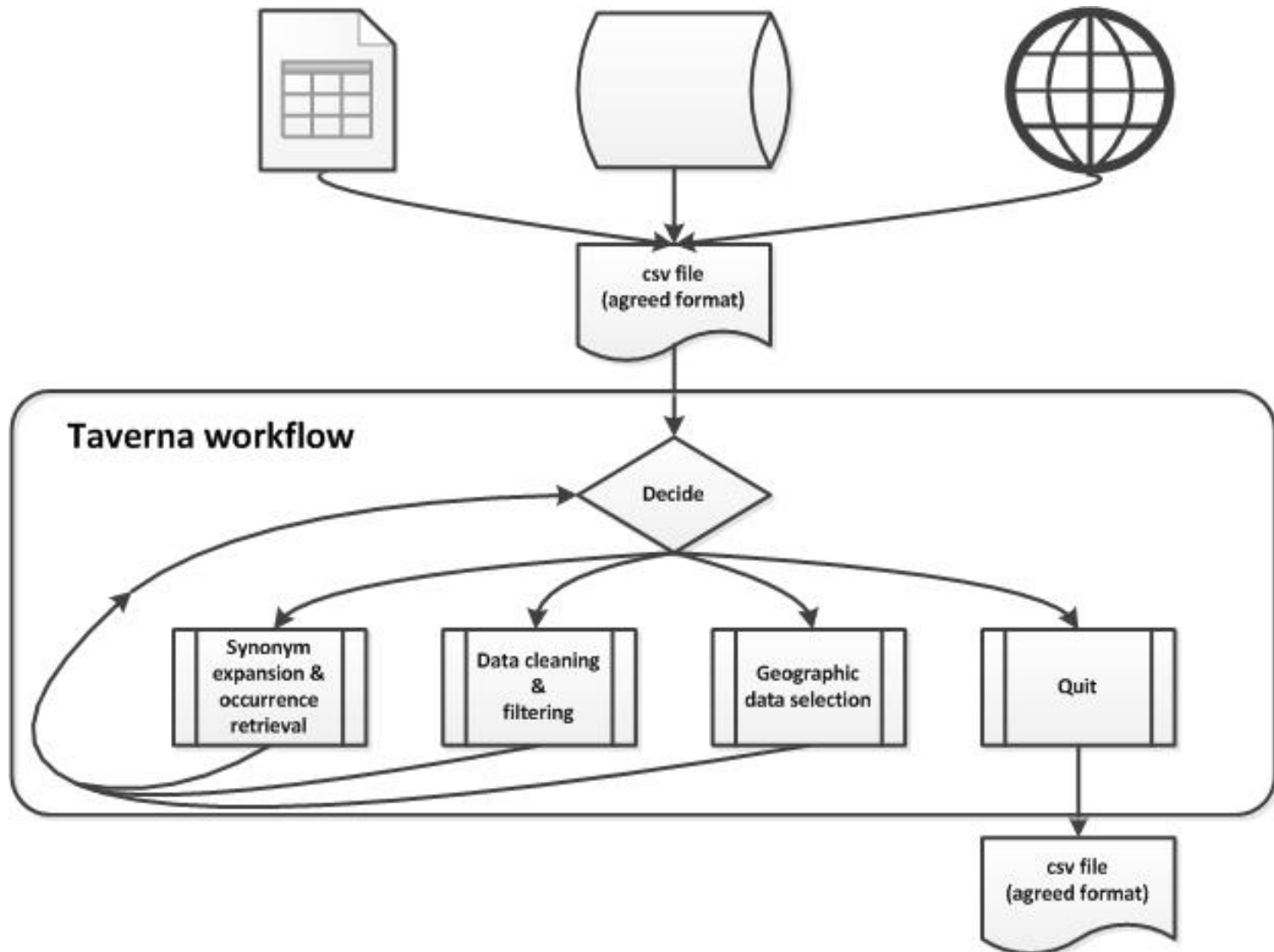
The name catalogue services (and related documentation) which correspond to Catalogue of Life data can be found at,

- [name search](http://dev.e-taxonomy.eu/cdmserver/col/name_catalogue) : http://dev.e-taxonomy.eu/cdmserver/col/name_catalogue
- [name information](http://dev.e-taxonomy.eu/cdmserver/col/name_catalogue/name) : http://dev.e-taxonomy.eu/cdmserver/col/name_catalogue/name
- [taxon information](http://dev.e-taxonomy.eu/cdmserver/col/name_catalogue/taxon) : http://dev.e-taxonomy.eu/cdmserver/col/name_catalogue/taxon
- [accepted name search](http://dev.e-taxonomy.eu/cdmserver/col/name_catalogue/accepted) : http://dev.e-taxonomy.eu/cdmserver/col/name_catalogue/accepted
- [fuzzy name search](http://dev.e-taxonomy.eu/cdmserver/col/name_catalogue/fuzzy) : http://dev.e-taxonomy.eu/cdmserver/col/name_catalogue/fuzzy

Data refinement workflow: integrated services

- Catalogue of Life (CoL)
- GBIF Checklist Bank
- GBIF name parser
- Swedish LifeWatch occurrence data (SWL)
- GBIF Occurrence Data
- Google Refine
- Biodiversity Spatial Temporal Interactive Interface (BioSTIF)
- infoXY (mapping coordinates to countries)

Data refinement workflow simplified



Data refinement workflow

Use occurrence data in a tabular format or just a list of latin names. Example:

nameComplete

Ameira divagans

Boccardia redeki

Bougainvillia rugosa

Branchiura sowerbyi

Cercopagis pengoi

Chelicorophium curvispinum

Chionoecetes opilio

Data refinement workflow


Running: Data Refinement Workflow v13

Action Required

Choose Sub-Workflow

Taxonomic Name Resolution / Occurrence Retrieval 

Taxonomic Name Resolution requires the input data file to contain a column named 'nameComplete', as explained [here](#). Other columns (if present) will not be considered for the purposes of the workflow run.

Data Quality (Google Refine) 

The possibilities for improving the quality of data are flexible and depend on the content of the input data file. As long as the file contains at least one named column of data the workflow will run.

Data Selection (BioSTIF) 

Geographical Data Selection (BioSTIF) requires the input data file to contain columns named 'decimalLatitude' and 'decimalLongitude' with valid coordinates, as explained [here](#). Do not select this option if your data file does not contain this information.

End Workflow

OK

Data refinement workflow

Running: Data Refinement Workflow v13

Action Required

Input Names

Ameira divagans

Cercopagis pengoi

Branchiura sowerbyi

Chelicorophium
curvispinum

Chionoecetes opilio

Boccardia redeki

Bougainvillia rugosa

Resolve Taxonomic Concept

Catalogue Of Life

[checklist_url](#)

Accepted Name



Chionoecetes opilio (O. Fabricius,
1788)

Synonym



Cancer opilio O. Fabricius, 1788



Chionoecetes behringianus Stimpson,
1857



Chionoecetes chilensis Streets, 1870



Peloplastus pallasi Gerstaecker, 1856

Retrieve (GBIF) species occurrence data

CANCEL

Data refinement workflow

Then continue the workflow with

- retrieving occurrence data from services such as GBIF,
- cleaning and enriching the data,
- selecting records needed for a specific analysis.

Please try yourself: **<https://portal1.at.biovel.eu/>**



Anton Güntsch

Freie Universität Berlin
Botanic Garden & Bot. Museum Berlin-Dahlem
Dept. Research and Biodiversity Informatics
Königin-Luise-Straße 6-8
D-14195 Berlin-Dahlem

a.guentsch@bgbm.org

