Checklistbank, use of different taxonomies and Catalogue of Life

Dave Martin

GBIF Global Biodiversity Information Facility

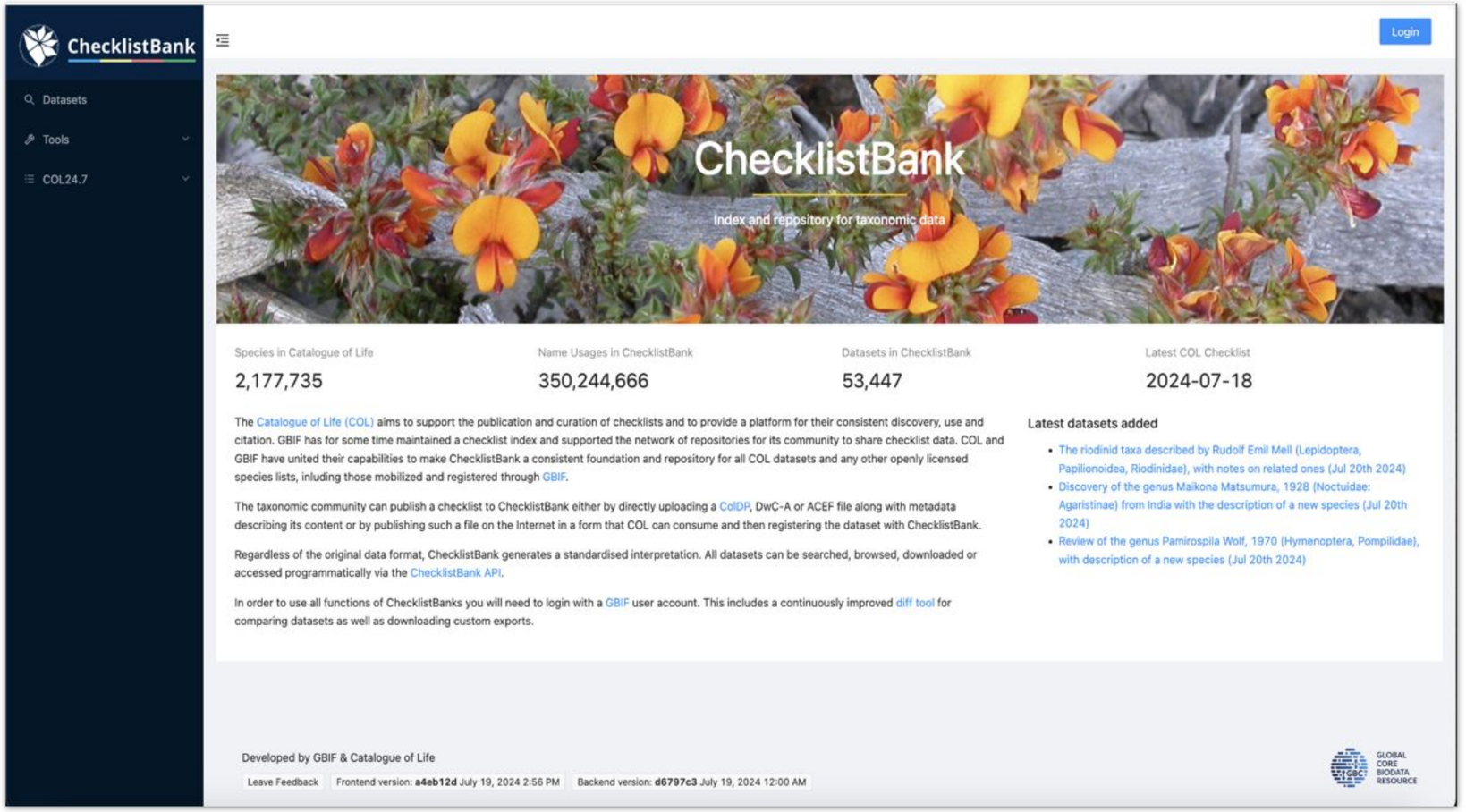
GBIF.org

Agenda

- ChecklistBank
- matching-ws
- Multi taxonomy support in GBIF API
- Processing as service

ChecklistBank

Taxonomic data infrastructure for everyone



Open data, open access taxonomic data publishing platform

Tools for:

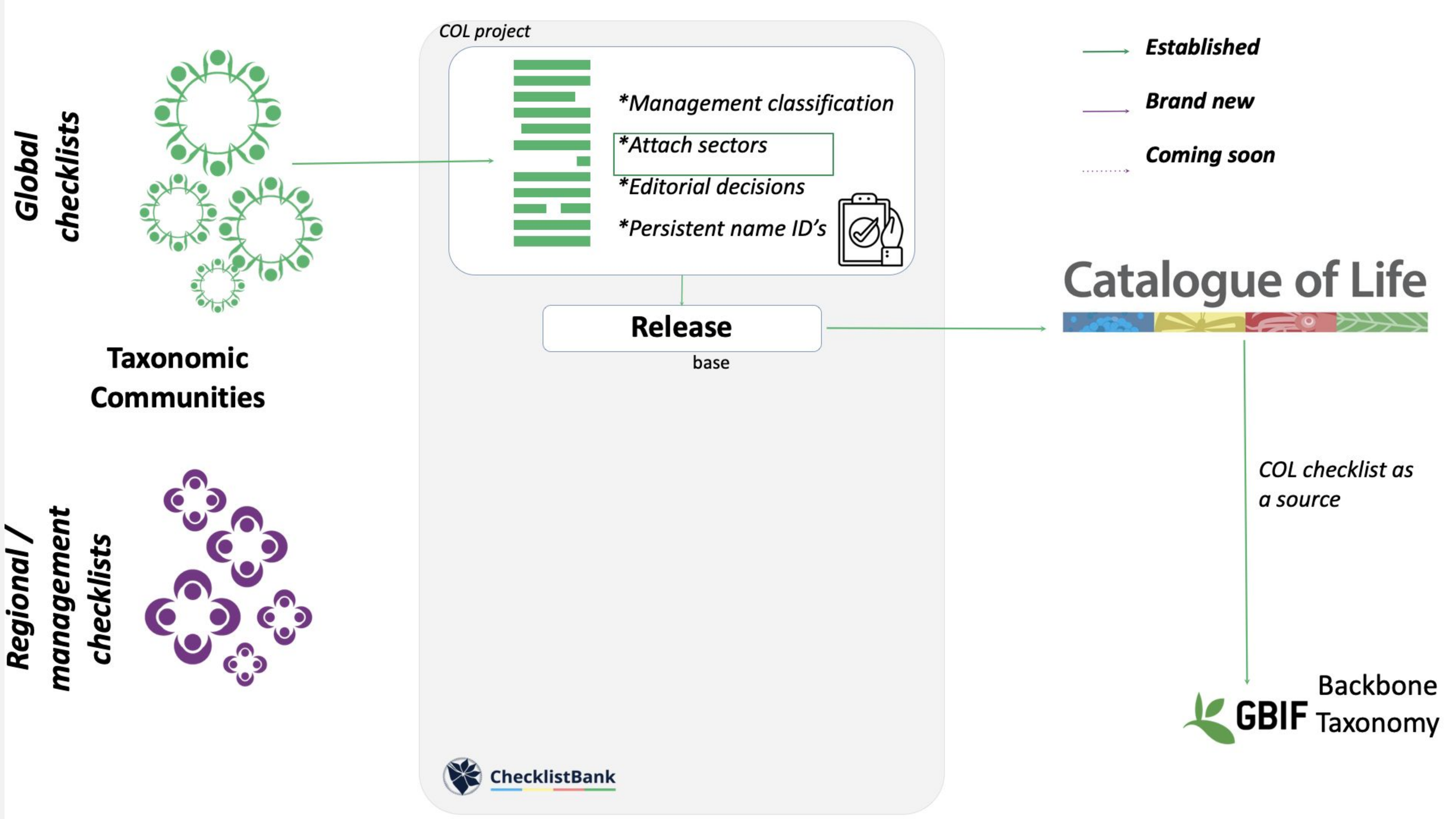
1. Importing lists
2. Comparing lists
3. Building lists

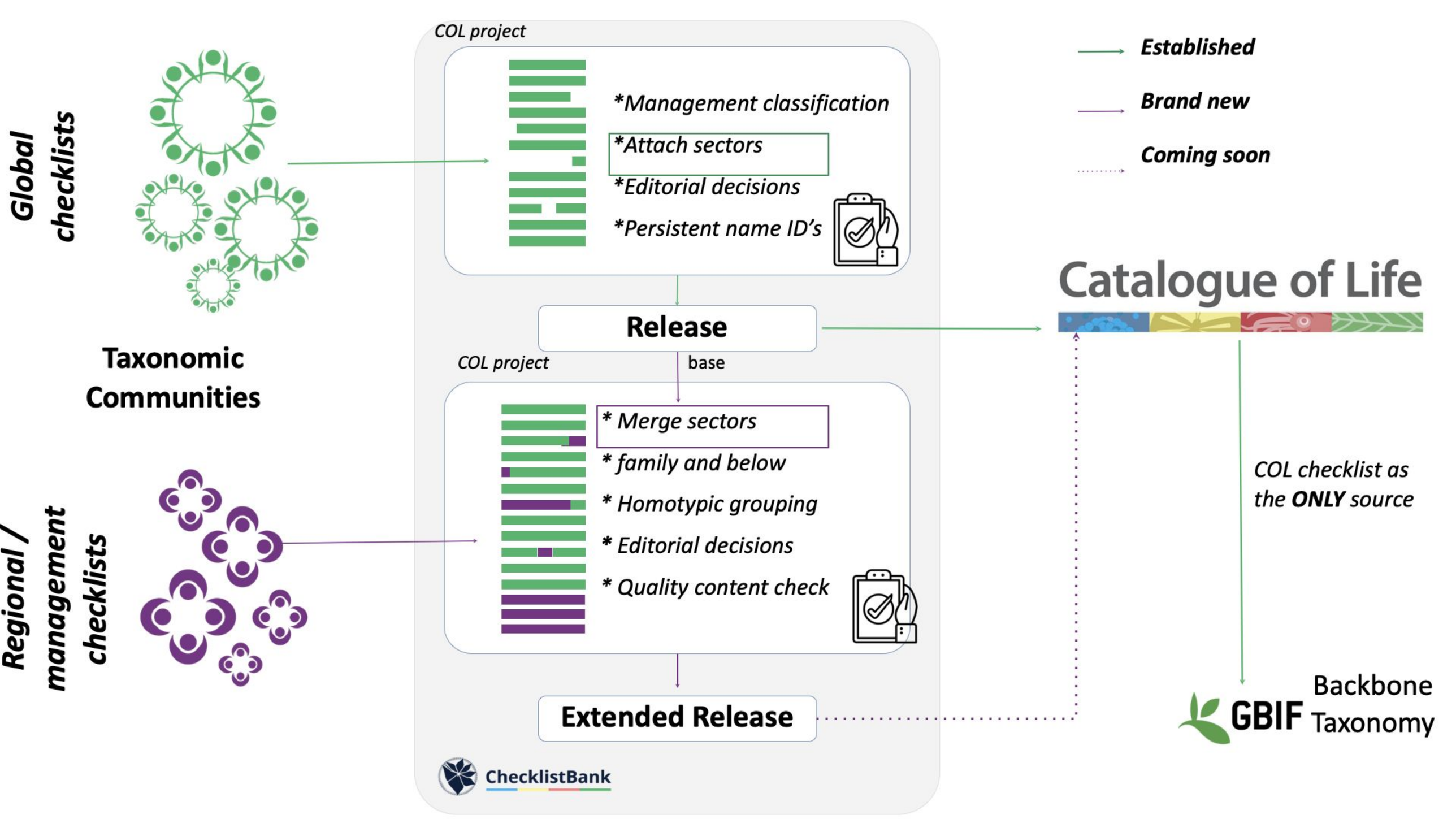
checklistbank.org

api.checklistbank.org

GBIF

Catalogue of Life

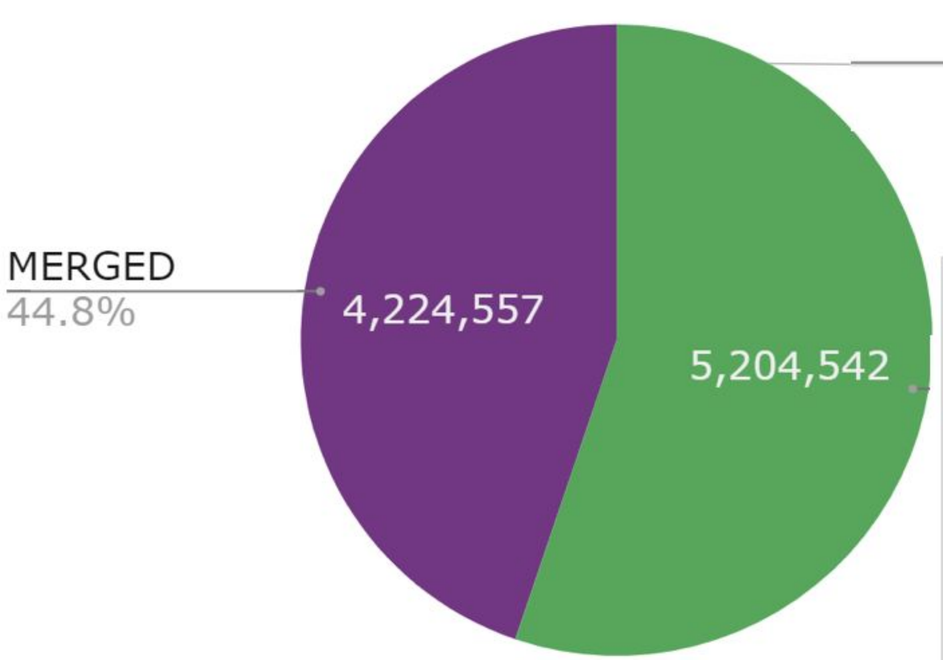

Global checklists
Taxonomic Communities
Regional / management checklists
COL project
*Management classification
*Attach sectors
*Editorial decisions
*Persistent name ID's
Release
base
ChecklistBank
Established
Brand new
Coming soon
Catalogue of Life
COL checklist as a source
GBIF
Backbone Taxonomy

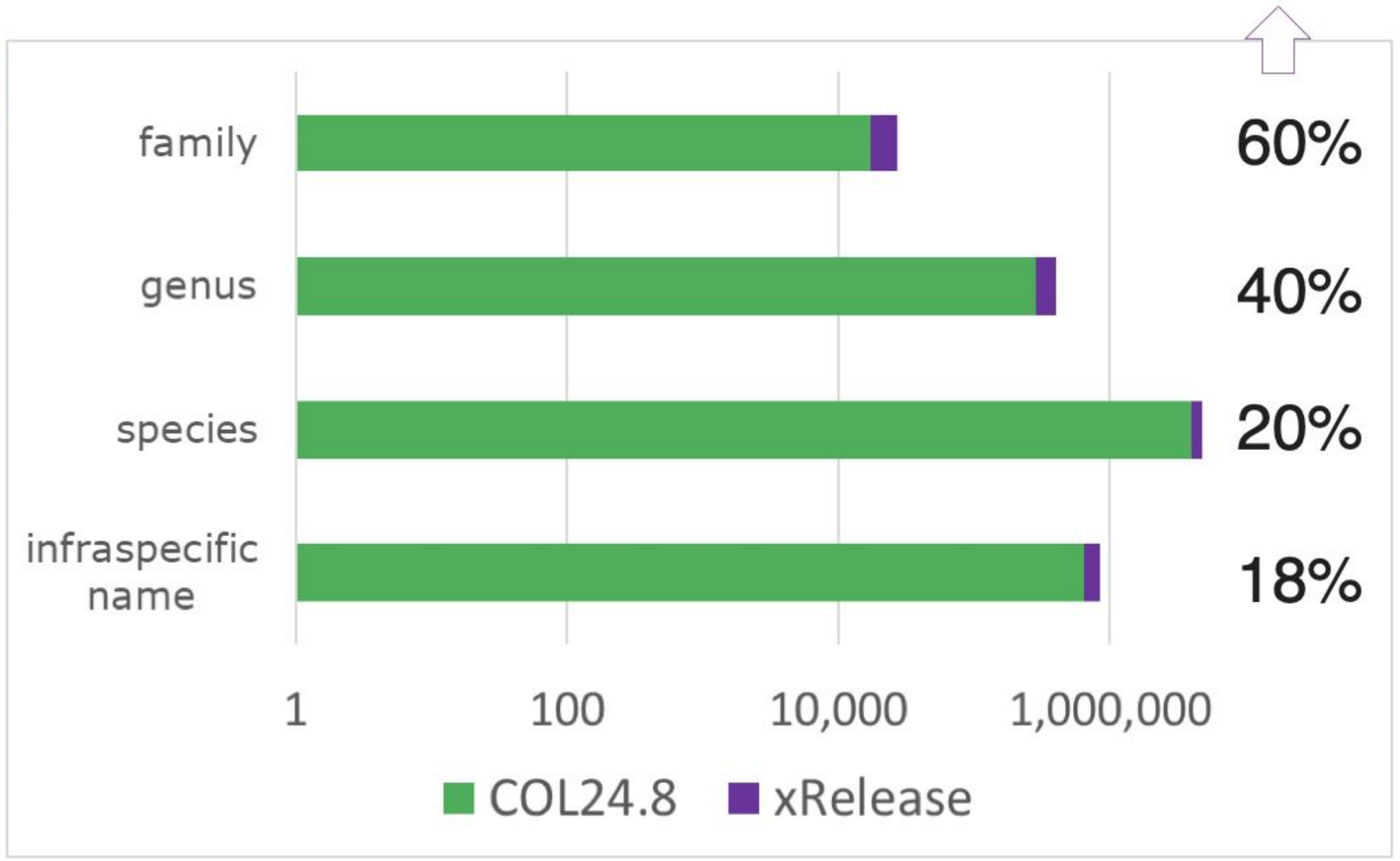

Global checklists
Taxonomic Communities
Regional / management checklists
COL project
*Management classification
*Attach sectors
*Editorial decisions
*Persistent name ID's
Release
COL project
base
* Merge sectors
* family and below
* Homotypic grouping
* Editorial decisions
* Quality content check
Extended Release
ChecklistBank
Established
Brand new
Coming soon
Catalogue of Life
COL checklist as the ONLY source
GBIF Backbone Taxonomy

Merged names in COL (all ranks)

xRelease
2024-08-25

Name count



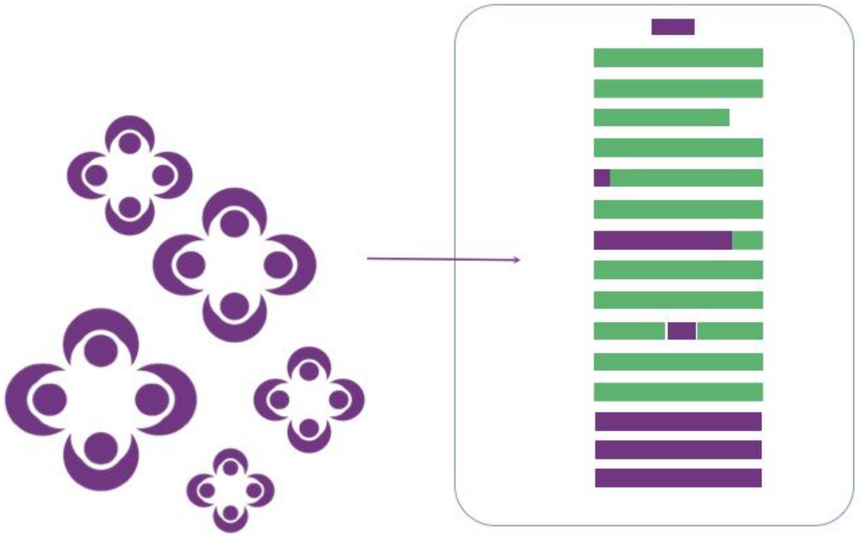


ONLY family names and

below are being merged

You can contribute to a better COL

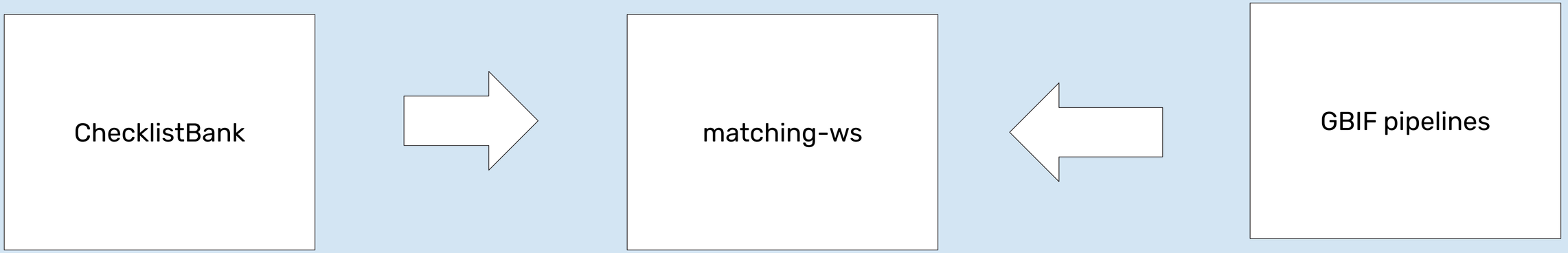
1. Publish or recommend a checklist



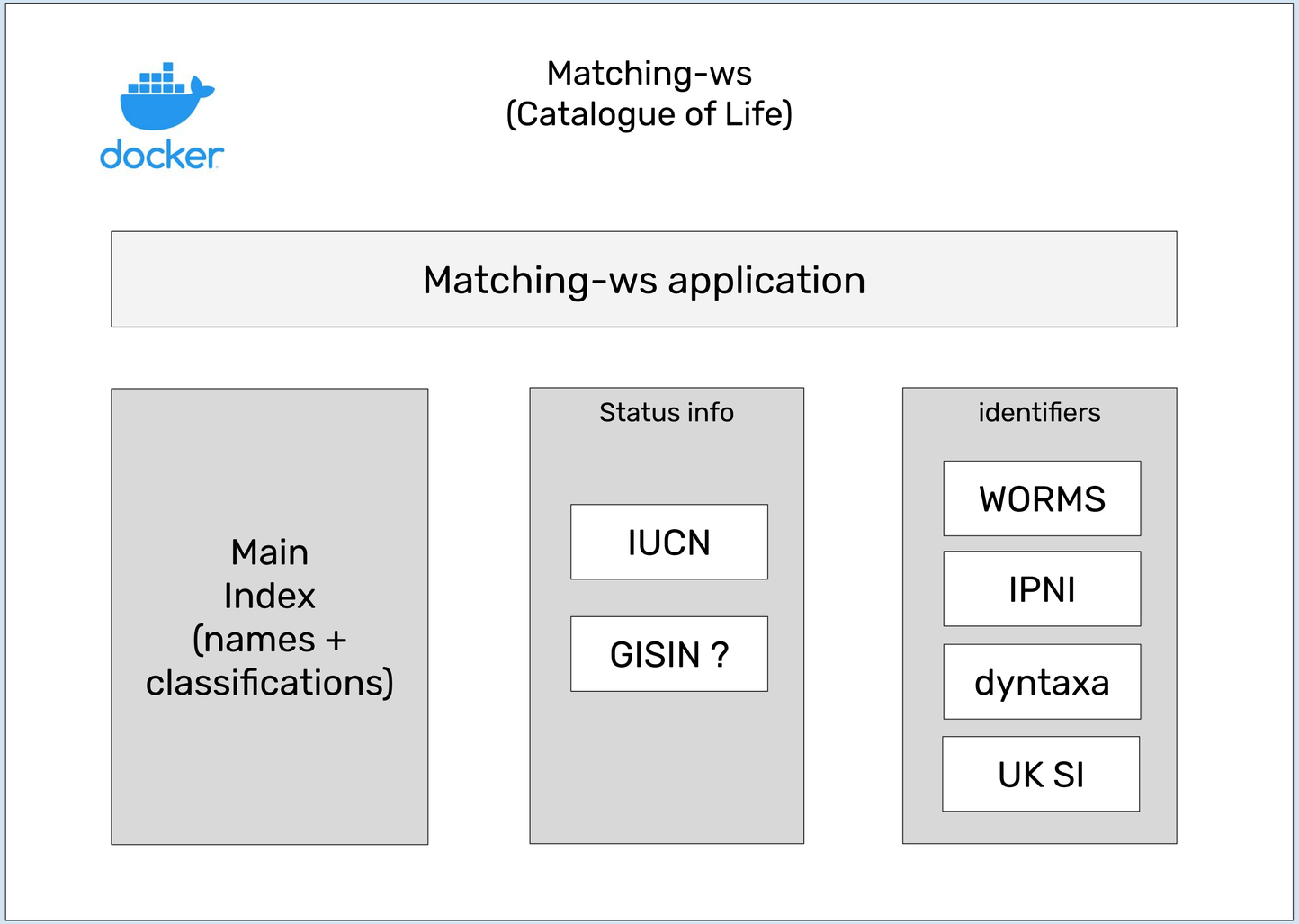
2. Issues report

https://github.com/CatalogueOfLife/data/issues

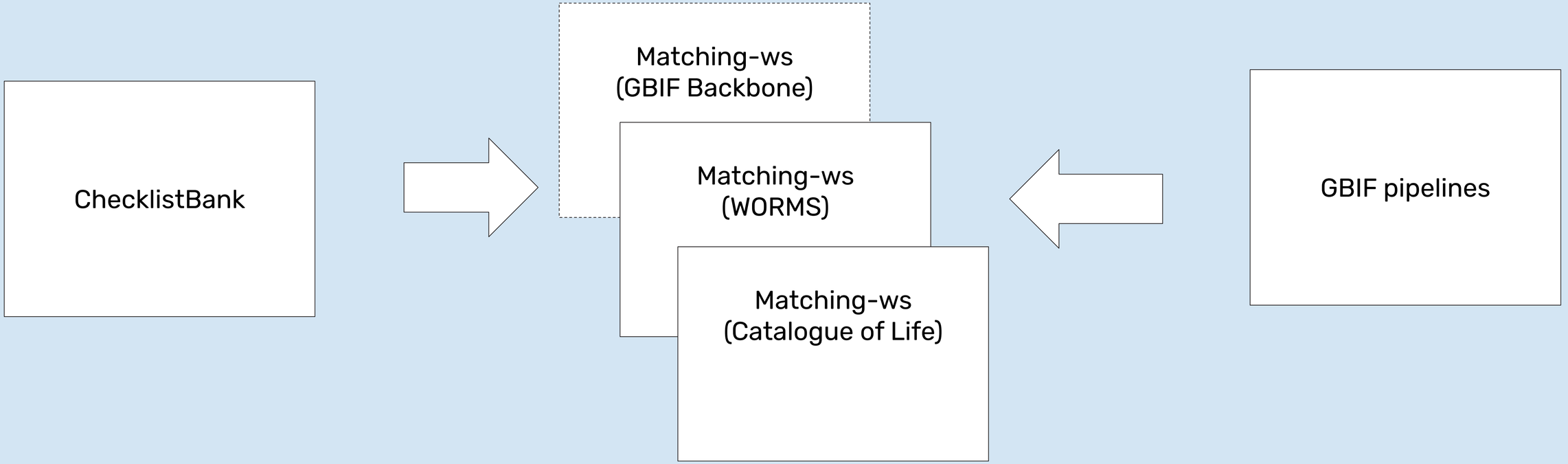
Matching-ws



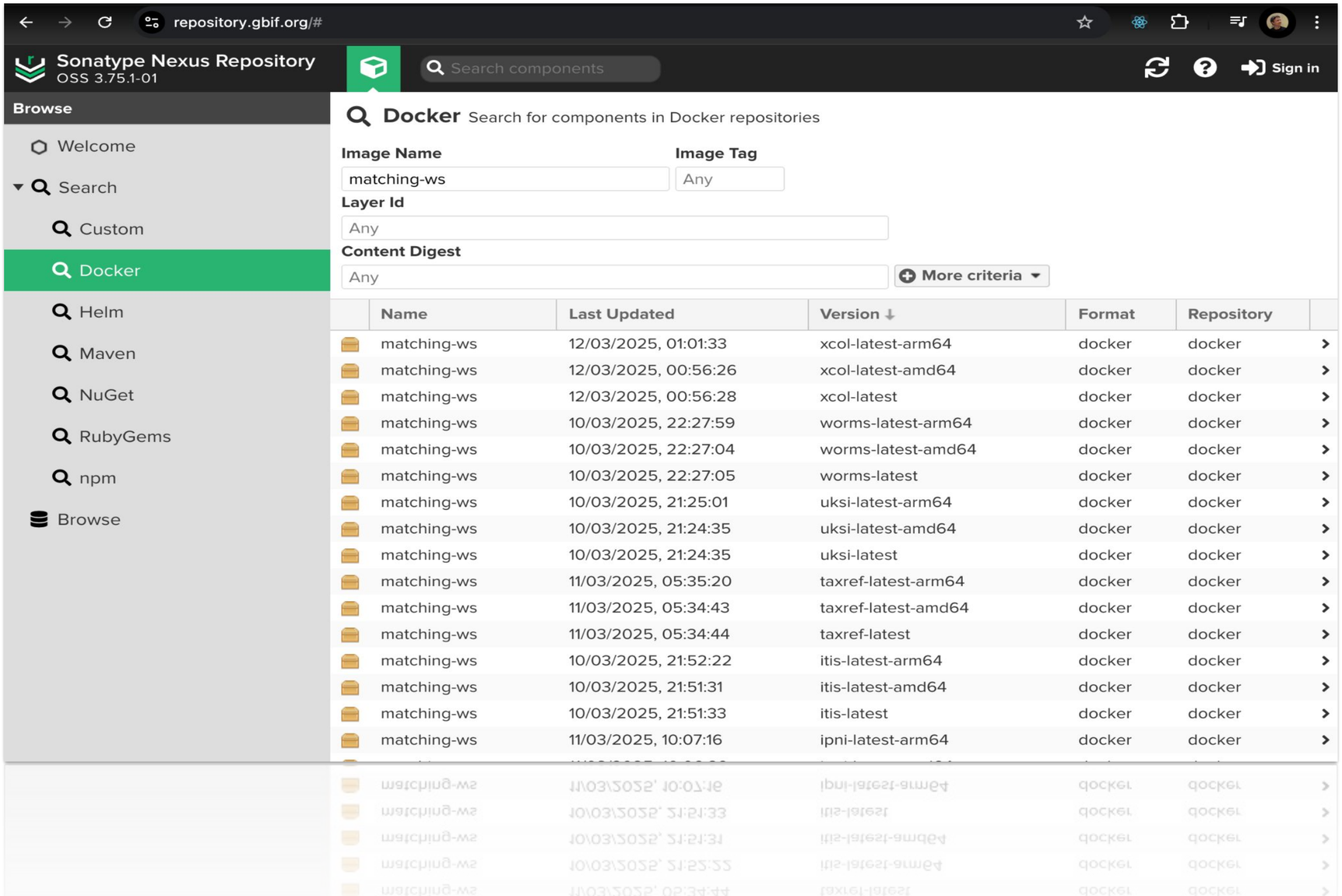
Matching-ws


docker
Matching-ws
(Catalogue of Life)
Matching-ws application
Main
Index
(names +
classifications)
Status info
IUCN
GISIN ?
identifiers
WORMS
IPNI
dyntaxa
UK SI

Matching-ws


ChecklistBank
Matching-ws
(GBIF Backbone)
Matching-ws
(WORMS)
Matching-ws
(Catalogue of Life)
GBIF pipelines

https://repository.gbif.org



Usage

```
docker run -p 8080:8080 docker.gbif.org/matching-ws:gbif-backbone-latest-amd64

docker run -p 8080:8080 docker.gbif.org/matching-ws:gbif-backbone-latest-arm64
```

Multi-taxonomy support in GBIF API

v2 Species match API

- `/v2/species/match?scientificName=Carcharodon+carcharias`

- `/v2/species/match?scientificName=Carcharodon+carcharias&checklistKey=2d59e5db-57ad-41ff-97d6-11f5fb264527`

WoRMS registry ID

v1 Occurrence API - search by name

- `/v1/occurrence/search?scientificName=Carcharodon+carcharias`

- `/v1/occurrence/search?scientificName=Carcharodon+carcharias&checklistKey=2d59e5db-57ad-41ff-97d6-11f5fb264527`

WoRMS registry ID

v1 Occurrence API - TaxonID

- `/v1/occurrence/search?taxonID=urn:lsid:marinespecies.org:taxname:225860&checklistKey=2d59e5db-57ad-41ff-97d6-11f5fb264527`

WoRMS registry ID

v1 Occurrence API - facets

- `/v1/occurrence/search?facet=familyKey`

- `/v1/occurrence/search?facet=subfamilyKey&checklistKey=7ddf754f-d193-4cc9-b351-99906754a03b`

CoL registry ID

v1 Occurrence download API

```
"predicate": {
    "type": "equals",
    "key": "TAXON_KEY",
    "value": "urn:lsid:marinespecies.org:taxname:225860",
    "checklistKey": "2d59e5db-57ad-41ff-97d6-11f5fb264527"
}
```

SQL API

```
SELECT

 decimalLatitude, decimalLongitude, eventDate

FROM occurrence

WHERE

array_contains(classifications["2d59e5db-57ad-41ff-97d6-11f5fb
264527"], "urn:lsid:marinespecies.org:taxname:158970")
```


registry ID

matching-ws

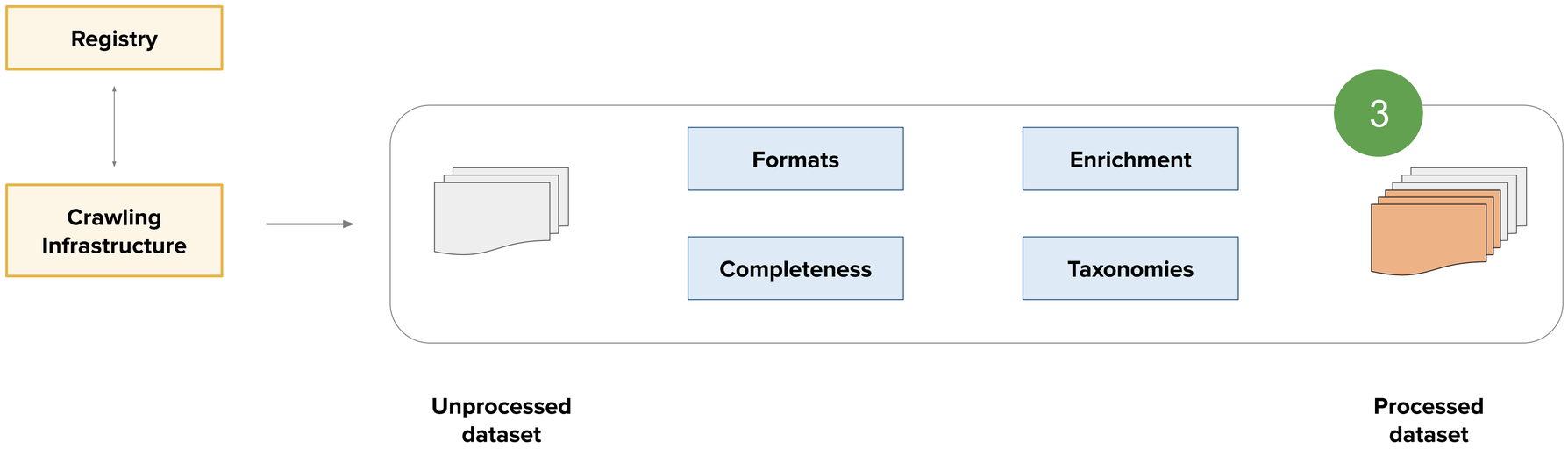
- “Catalogue of Life ready”
  - String identifiers
  - Support for intermediary ranks (eg. subphylum, subfamily)

- Consolidated ID support
  - WoRMS, IPNI, Dyntaxa, UK NBN

- Includes IUCN status
  - Scope for additional status values (e.g. invasive)

Processing as a service

Idea: Processing as a service?



GBIF Data Warehouse (simplified)

Source Systems

Relational Databases

Triple stores

Object stores

Spreadsheets

APIs

1

Registry

Crawling Infrastructure

Data pipelines

Formats

Consistency

Completeness

Enrichment

Taxonomies

2

APIs

Reference Catalogues

3

Indexes

SQL services

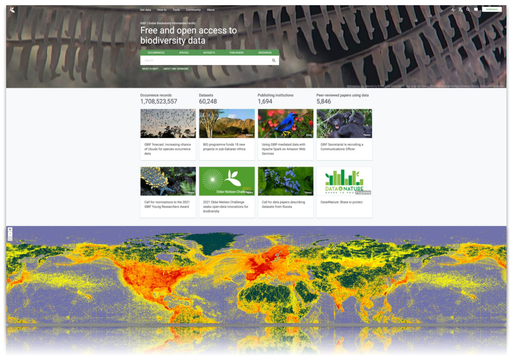
Reports

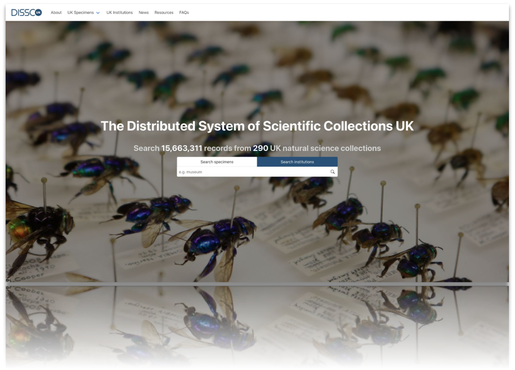
Image caches

Open APIs

Citation tracking

4





Thanks!