

## WP2 - Work package description

|                         |   |
|-------------------------|---|
| Work package number     | WP2   |
| Type of activity        | COORD   |
| Work package title      | Integration of the Virtual Research Community |
| Start month             | 1   |
| End month               | 36  |
| Lead beneficiary number | 1   |

### Objectives:

To establish and integrate a Virtual Research Community that will enhance and unify taxonomic enumeration and recording of the global biota among the major global biodiversity programmes, building on the Catalogue of Life as a common platform. This will involve:

#### 1) Coordination of the Virtual Research Community

A series of workshops and discussions to plan workflows and protocols between the organisations, as needed to share knowledge and to unify the process of extending the taxonomic knowledge backbone. In particular these will include a series of design and specification workshops, and testing sessions for the following broad workflow areas:

- establishing a workflow that delivers refreshed instances of Catalogue of Life taxonomy within the very heterogeneous working platforms and portals of the global programme partners (GBIF, EBI, ECBOL, IUCN and LifeWatch), using a CoL download service as its base.
- establishing workflows for some of the global programme partners, that enables them to cross-map their own taxonomies to the CoL taxonomy, and to utilise their own and the CoL side-by-side within their own system using a CoL Cross-mapping Service developed in WP11.
- establishing workflows that enable some of the global programme partners that maintain their own taxonomies or name stores to contribute records for research and taxonomic positioning by the CoL array of taxonomic expert organisations, using a CoL GSD-Piping Service developed in WP12. This part will use a CoL Reception Buffer that will either contain or be linked to a Name Index that will use the Name Index structure under discussion in the Global Names Architecture (GNA) group and worked on by partners in EoL and GBIF.

#### 2) Coordination of Tools and Services

The co-ordination of 'central' CoL tools and services needed for these workflows (eg in WP3, WP4, WP11, and WP12)

### 3) Implementations in the Global Partner Organisations

The co-ordination of implementations of the workflows in the global programme partner organisations (eg in WP6, WP7, WP8, WP9 and WP10).

### 4) Workshops

one or more workshops to address policy and implementation issues on some of the scientific challenges to sharing a common taxonomy between these communities, for example: the global names architecture (GNA); clusters of sequences detected in environmental metagenomics, cryptic species detected in barcoding, and infra-specific taxa of importance in conservation.

## Description of work and role of partners:

Many of the partners have scoping, design, test and acceptance responsibilities in this workpackage for particular tools or workflows for which they are responsible for the actual development in small separate workpackages under Joint Research Activities: WP6, WP7, WP8, WP9, WP10, WP11, WP12.

#### Task 2.1

Coordination of the Virtual Research Community Scientific organisation of the workshops and co-ordinated tasks of WP2, co-ordination with the global programme partners work (WP7, WP8, WP9, WP10), and preparation of new software(WP4, WP11 & WP12).

#### Task 2.2

Coordination of CoL Download Tools Co-ordinate establishment of workflows that deliver and refresh instances of the CoL in platforms of the major programmes, including a) specification, b) test and possible improvement, and c) acceptance.

#### Task 2.3

Coordination of CoL Cross-mapping Tools Co-ordinate establishment of cross-mapping workflows, including a) specification, b) test and possible improvement, and c) acceptance.

#### Task 2.4

Coordination of CoL GSD Piping Tools Co-ordinate establishment of workflows that deliver batches of name- and taxon-records to the GSD-Piping Service via the CoL Reception Buffer and taxonomic placement by the Sp2000 Taxonomic GSD and Expert Array, including a) specification, b) test and possible improvement, and c) acceptance.

#### Task 2.5

Implementations in the Global Partner Organisations Coordinate implementations in Global Partner Organisations

## Task 2.6

Workshops Convene workshops on taxonomic policy and implementation issues (eg global names architecture indexes (GNA), environmental metagenomics BINS & FESINS, cryptic species and barcoding).

### Person-Months per Participant:

| Participant number | Participant short name | Person-months per participant |
|--------------------|------------------------|-------------------------------|
| 1                  | UR                     | 51.20                         |
| 2                  | GBIFS                  | 9.00                          |
| 3                  | EMBL                   | 9.00                          |
| 4                  | USNM                   | 1.00                          |
| 5                  | KNAW                   | 4.00                          |
| 6                  | MIZ-PAS                | 5.00                          |
| 7                  | FUB-BGBM               | 9.00                          |
| 8                  | UICN                   | 9.00                          |
| 10                 | ETI Bioinformatics     | 4.00                          |
| 11                 | CU                     | 11.00                         |
| 12                 | MNHN                   | 16.60                         |
|                    | Total                  | 128.80                        |

### List of deliverables:

| Deliverable Number | Deliverable Title                        | Lead Beneficiary number | Estimated indicative person-months | Nature | Dissemination level | Delivery date |
|--------------------|--|-------------------------|------------------------------------|--------|---------------------|---------------|
| D2.1               | Download and Piping Tools Specifications | 1                       | 21.90                              | R      | RE                  | 6             |
| D2.2               | Enhanced Cross-map Service Specification | 1                       | 18.60                              | R      | PU                  | 8             |
| D2.3               | Workshop 1 Policy Report                 | 1                       | 6.40                               | R      | PU                  | 12            |
| D2.4               | Initial Implementations                  | 1                       | 6.00                               | R      | PU                  | 22            |
| D2.5               | Workshop 2 Policy Report                 | 1                       | 5.30                               | R      | PU                  | 24            |
| D2.6               | Workshop 3 Policy Report                 | 1                       | 5.30                               | R      | PU                  | 24            |
| D2.7               | Taxonomic Integration                    | 1                       | 65.30                              | R      | PU                  | 36            |
|                    | Total                                    |                         | 128.80                             |        |                     |               |

## Description of deliverables:

D2.1)

Download and Piping Tools Specifications: Specification for the CoL Download and GSD Piping Tools, as agreed by WP2.

[month 6]

D2.2)

Enhanced Cross-map Service Specification: Specification for the Enhanced Cross-map Service, as agreed by WP2.

[month 8]

D2.3)

Workshop 1 Policy Report: Policy Report, possibly relating to the Global Names Architecture Plan. [month 12]

D2.4)

Initial Implementations: Report on the initial implementations of the CoL service pipelines at the Global Partner sites.

[month 22]

D2.5)

Workshop 2 Policy Report: Policy Report, possibly relating to classification and naming of DNA clusters in Environmental Genomics.

[month 24]

D2.6)

Workshop 3 Policy Report: Policy Report, possibly relating to taxonomic handling of cryptic variation and the results of bar-coding.

[month 24]

D2.7)

Taxonomic Integration: Integration of combined species catalogues across the Global Organisations.

[month 36]

**Schedule of relevant Milestones:**

| Milestone number | Milestone name  | Lead Beneficiary number | Delivery date from Annex I | Comments         |
|------------------|---|-------------------------|----------------------------|------------------|
| MS1              | CoL Enhanced Download Service specification available | 1                       | 2                          |                  |
| MS2              | Preliminary Cross-map Service Specification available | 1                       | 4                          |                  |
| MS3              | CoL Enhanced Download Service in test                 | 10                      | 5                          | Provided by WP4  |
| MS7              | Preliminary Cross-mapping Tools in test               | 11                      | 8                          | Provided by WP11 |
| MS10             | Taxon & Name Supply-side Pipeline in test             | 1                       | 12                         | Provided by WP12 |
| MS15             | Enhanced Cross-mapping Tools in test                  | 11                      | 14                         | Provided by WP11 |
| MS18             | Taxon & Name Distribution-side Pipeline in Test       | 1                       | 16                         | Provided by WP12 |