SYNTHESYS+
Synthesis of Systematic Resources
a DiSSCo project

D1.13 DATA MANAGEMENT PLAN

Katherine Dixey, Matt Woodburn, Vincent Smith, Quentin Groom, Wouter Addink

Grant Agreement Number | 823827
Acronym | SYNTHESYS PLUS
Call | H2020-INFRAIA-2018-2020
Start date | 01/02/2019
Duration | 48 months
Work Package | WP1 NA1
Work Package Lead | Vince Smith
Delivery date | 10/09/2019

SYNTHESYS+ was funded by the Horizon 2020 Framework of the European Union under the H2020 Open Innovation and Open Science Research Infrastructure call.

SYNTHESYS+
Synthesis of Systematic Resources
a DiSSCo project

Contents

SYNTHESYS+ was funded by the Horizon 2020 Framework of the European Union under the H2020 Open Innovation and Open Science Research Infrastructure call.

Project Summary

European natural history collections are a critical infrastructure for meeting the most important challenge humans face over the next 30 years – mapping a sustainable future for ourselves and the natural systems on which we depend – and for answering fundamental scientific questions about ecological, evolutionary, and geological processes. Since 2004 SYNTHESYS has been an essential instrument supporting this community, underpinning new ways to access and exploit collections, harmonising policy and providing significant new insights for thousands of researchers, while fostering the development of new approaches to face urgent societal challenges. SYNTHESYS+ is a fourth iteration of this programme and represents a step change in evolution of this community. For the first time SYNTHESYS+ brings together the European branches of the global natural science organisations (GBIF, TDWG, GGBN and CETAF) with an unprecedented number of collections, to integrate, innovate and internationalise our efforts within the global scientific collections community. Major new developments addressed by SYNTHESYS+ include the delivery of a new virtual access programme, providing digitisation on demand services to a significantly expanded user community; the construction of a European Loans and Visits System (ELViS) providing, for the first time, a unified gateway to accessing digital, physical and molecular collections; and a new data processing platform (the Specimen Data Refinery), applying cutting edge artificial intelligence to dramatically speed up the digital mobilisation of natural history collections. The activities of SYNTHESYS+ form a critical dependency for DiSSCo - the Distributed System of Scientific Collections, which is the European collection communities ESFRI initiative. DiSSCo will undertake the maintenance and sustainability of SYNTHESYS+ products at the end of the programme.

SYNTHESYS+ Data Management: Background and Principles

This data management plan is based on the Horizon 2020 DMP template. The SYNTHESYS+ DMP will be updated once per year whilst the project is active, to ensure the content remains relevant and is as comprehensive as possible.

As SYNTHESYS+ is a DiSSCo-linked project, the DMP will align with the scientific vision and mission of the DiSSCo RI and the Provisional DiSSCo DMP (under construction in ICEDIG WP6.2). The DiSSCo DMP describes the main DiSSCo data management principles and requirements. The DiSSCo DMP will be made publicly available on the ICEDIG project website (www.icedig.eu) once the first version is complete (est. October 2019).

Schedule of Updates

V1 submitted 10/09/2019

V2 reviewed by 31/08/2020

V3 reviewed by 31/08/2021

V4 reviewed by 31/08/2022

SYNTHESYS+ was funded by the Horizon 2020 Framework of the European Union under the H2020 Open Innovation and Open Science Research Infrastructure call.

1. Data Summary

This section will summarise the following information for each work package:

- The purpose of the data collection/generation and its relation to the project objectives
- The types and formats of data the project will generate and collect
- Details of any data reuse
- The origin of the data
- The expected size of the data (where applicable)

A summary of data collected under each work package, along with applicable licensing and policies, can be seen in the table below:

VIEW DATA SUMMARY TABLE

Work Package Data Summaries

Networking Activities

WP2 | NA2 - Harmonisation of policies and best practices, training and support

NA2 will ensure effective harmonisation, delivery and dissemination of best practices and standard operating procedures across the consortium, assisting in training, providing helpdesk support and developing support resources as required. The key issues in this work package relate to coordination in policy development processes, harmonisation of policy implementation actions and the support activities, including training and helpdesk. The Distributed European School of Taxonomy (DEST) will be utilised for training actions when feasible.

Expected data types and data re-use: The ticketing system for the helpdesk developed under **Task 2.4** will be embedded into ELViS and may entail collection of limited amounts of personal data, which will be handled within the same system as data collected under **TA1**. The unified collections data dashboard under **Task 2.2** will re-use existing data from sources such as CETAF passports, the SYNTHESYS Collections Self-Assessment Tool, and One World Collection assessments. This dashboard will be embedded into the European Loans and Visits System (ELViS) - more information on data types and usage within ELViS can be found under **JRA1.** For **Task 2.1** a range of policy documents will be analysed, and datasets from policy assessments under ICEDIG may also be re-used. Data standards will be collected in collaboration with work under **NA3** and **NA4**. This task will result in a published policy implementation manual which will be openly published online for use by the community. **Task 2.3** will generate a catalogue of training modules on subjects such as data processing, data standards, policy and legislation, DNA barcoding, and collections management software, which will be made openly available on online training platforms wherever feasible.

WP3 | NA3 - Molecular collections in the age of genomics - standards & processes

Work under NA3 will be developed in close association with the activities of GGBN (www.ggbn.org). Tasks include:

- Identification of institutions that house a biodiversity biobank and are willing to operate access to their molecular collections (**Task 3.1)**

SYNTHESYS+ was funded by the Horizon 2020 Framework of the European Union under the H2020 Open Innovation and Open Science Research Infrastructure call.

SYNTHESYS+ Synthesis of Systematic Resources a DiSSCo project

- Development of a certification system of biodiversity biobanking facilities associated with NH collections (**Task 3.2**)
- Development of mechanisms that secure free movements of data and, when possible, samples between these certified institutions (**Task 3.3**)

To reach these goals a landscape analysis of biobank standards will be completed and best practices defined for use of molecular collections.

Expected data types and data re-use: This work package will produce a set of documented molecular standards and a landscape analysis report (Deliverable 3.1), and a handbook detailing best practice for usage of molecular collections (Deliverable 3.2). The planned registration system for biodiversity biobanks (Deliverable 3.3) will create a system that facilitates biobanking data integration with the GGBN portal - this will entail the processing of institutional information and biobank sample records.

WP4 | NA4 - Digital standards & processes

NA4 will develop standards which facilitate technical coordination between institutions, through data linking and improved interoperability. **Task 4.1** will support standards development across SYNTHESYS. This will be done be helping partners navigate the standards development process from the launching of task groups to the ratification process. The task will provide open tools for testing standards, such as Wikibase, but also guidance for how standards should be documented. The standards to be worked on will be those needed by the community, particularly those to achieve DiSSCo. The task will also ensure that standards development in Europe is integrated in international standards development. Finally, the work of this task will be documented in a landscape analysis that will point to the remaining priority areas for standards development to achieve DiSSCo. **Task 4.2** will increase community adoption of the stable identifier framework for specimens developed under CETAF-ISTC. **Task 4.3** will work to develop a set of API specifications for the International Image Interoperability Framework (IIIF) to ensure interoperability of image servers across European institutions.

Expected data types and data re-use: The primary output of NA4 will be a range of documentation and standards as outlined above. Furthermore, a monitoring dashboard will be developed to show institutional adoption and compliance within Europe - this will utilise data collected under tasks 4.1 and 4.3.

Work completed under NA4 will influence project and wider community data management practices and will play a key part in the future development of this document.

WP5 | NA5 - Internationalisation and engagement of new user communities in EC priority areas

NA5 is designed to facilitate coordination between the SYNTHESYS community and a range of relevant stakeholders through facilitation of workshops and meetings. These will cover: development of international roadmaps for biodiversity infrastructure (**5.1**); geographic expansion of the user base (**5.2**); and alignment and coordination with EU stakeholders (**5.3**).

Expected data types and data re-use: Though largely networking-oriented the work package will generate a series of engagements reports, a roadmap, and a whitepaper.

Joint Research Activities

WP6 | JRA1 – Optimisation of Access

The European Loans and Visits System (ELViS) will provide open access to over 490 million specimens at 21 SYNTHESYS+ institutions. This will be utilised by the SYNTHESYS+ Transnational Access and Virtual Access (Digitisation on Demand) programme. The system will collect and generate data to facilitate the placement, assessment, prioritisation and monitoring of requests for visits, loans and requests for digitisation on demand (DoD). The system will continue as a DiSSCo eService after the project.

Expected data types and data re-use: The principal data types are institutional information, facility and equipment data, institutional staff information, researcher profile and expertise, transactions including communications for loans, visits and DoD. To provide discoverability of collections, ELViS will also collect data about the collections in the facilities to create a proto-European Collection Objects Index. The data will include PIDS, textual data & metadata, and links to data in external trusted repositories including text and images. Data will initially be stored in spreadsheets and in a database (for the first Virtual Access call with DoD requests) and later for development of ELViS in the form of Digital Objects (DO) in CORDRA or a similar DO repository. A wide range of pre-existing external data, such as institutional information, facility and equipment details, and researcher profiles will be made available to ELViS and other DiSSCo services from multiple external authoritative sources (GBIF, CoL, Plazi, others). Collections data will be sourced from facilities, usually from Collections Management Systems. For digital specimen data, ELViS will use aggregated and quality controlled data from GBIF as trusted repository.

Data size: Since ELViS will link to data in external repositories, the data stored centrally is expected to be limited to hundreds of megabytes or a few gigabytes.

WP7 | JRA2 - Collections on Demand

Task 7.1 will focus on the co-ordination of Virtual Access (VA) calls, which will be run through ELViS. Further relevant details can be found under **VA** and **JRA1.** Task 7.2 aims to establish Digitisation on Demand (DoD) workflows for collections, focusing on 3D images.**Task 7.3** will develop workflows, protocols and processes for Sequencing on Demand (DNAoD) and will therefore entail generation of molecular data. Work on DNAoD will be carried out in close alignment with molecular standards activities under **NA3**.

Expected data types and data re-use: For data relating to Task 7.1, see VA & JRA1. Task 7.2 will generate MicroCT datasets (TIFF/BMP/JPEG/PNG) alongside 3D analysis (.csv/text files) and surface models (.obj format). For Existing MicroCT datasets held by participants will be made available to end-users, and current datasets may be used for comparison purposes. Task 7.3 will generate new molecular data and may utilise existing datasets. Data may be subject to Access & Benefit Sharing (ABS) regulations.

WP8 | JRA3 - Specimen Data Refinery

JRA3 will develop a “Specimen Data Refinery” (SDR) platform that will integrate artificial intelligence with human-in-the-loop processes to extract, enhance and annotate data from digital images and high volumes of specimen records. **Task 8.1** will consist of a landscape analysis, evaluating existing platforms, approaches, services and systems and resulting in a written report. **Task 8.2** entails a

SYNTHESYS+ was funded by the Horizon 2020 Framework of the European Union under the H2020 Open Innovation and Open Science Research Infrastructure call.

series of subtasks which will each aim to develop an SDR service (e.g. optical character recognition; specimen image analysis). **Task 8.3** will develop the SDR Execution platform which will handle execution of the workflows and the user interface. Finally, **Task 8.4** will produce a summary report outlining use cases for delivery and data exploitation.

Expected data types and data re-use: JRA3 will seek to reuse openly available online specimen data from GBIF and institutional data portals. This work package will generate and collect: specimen data and images, process and configuration metadata, machine learning models, software code, user stories, and general project data.

Data size: The total size of the data generated and re-used is not expected to exceed 1TB in the first year.

Access

WP9 | TA1 - Transnational Access

The Transnational Access (TA) programme allows scientists and other potential users of natural history collections to apply for funding for short research visits at 21 participating institutions. The TA programme is currently run through an online application system, which collects and stores both application data and applicant information. A limited amount of institutional information, such as available facilities, is also stored. Management of this data is described in full in the following document:

Procedures for collection, use, protection and retention of personal data collected under the Transnational Access (TA) programme

Users of the TA programme are subject to institutional data management policies for data generated during their visit. During the first year of SYNTHESYS+, work will be done to create a unified set of TA data management guidelines that can be flexibly applied to all participating institutions, to ensure that SYNTHESYS-generated data is treated in accordance with FAIR principles.

All users of the TA programme are asked to provide details of anticipated publications, theses, datasets or other outputs arising from their visit as a requirement of their funding. Details of these are stored on the application site and can be updated with links at any time. Management of TA research outputs will eventually move to ELViS - see information under **JRA1** for further details.

WP10 | VA1 - Virtual Access

The new Virtual Access (VA) programme will prioritise community-driven requests for digitisation of specimens held within participating institutions. We can anticipate that this may include any or all of the following data types:

- Specimen data
- Specimen images and datasets
- Specimen analysis data
- Molecular data

SYNTHESYS+ was funded by the Horizon 2020 Framework of the European Union under the H2020 Open Innovation and Open Science Research Infrastructure call.

Until the first set of digitisation requests have been received, it is not possible to accurately predict the size of the expected data. This document will be updated following the close of the first VA Call in Spring 2020.

It is a fundamental stipulation of VA that all resulting data will be made openly accessible for use by all. This will be done through institutional data portals and aggregators (e.g. GBIF). There will be no embargoes imposed on Virtual Access data.

Data Utility

Data generated across all SYNTHESYS+ work packages will serve a variety of stakeholders both within and outside of the natural history collections community. This may include:

- Researchers engaged in discovering, describing and interpreting life on Earth, both past and present, as well as researchers studying the geological history of the planet.
- Citizen scientists, artists and students who need access to the natural collections.
- Curators and Collection Managers in charge of planning loans and visits.
- Digitisation staff and lab facilities staff.
- Institutional managers and directors of the collection facilities.
- Funders of collections and digitisation projects.
- SYNTHESYS+ project stakeholders

Publications & Research Outputs

Publications

In adherence to article 29.2 of the grant agreement, all beneficiaries will ensure open access to all peer-reviewed scientific publications. Before or on publication, a machine-readable electronic copy of the published version or final peer-reviewed manuscript will be placed in a repository for scientific publications, and any research data needed to validate the results will also be deposited. If permissible by the publisher, open access will be ensured on publication, or else within six months of publication.

Bibliographic metadata of scientific publications will include:

- The terms “European Union (EU)” and “Horizon 2020”
- The name of the action H2020-INFRAIA-2018-1, SYNTHESYS PLUS, 823827
- The publication date and any applicable embargoes
- A persistent identifier

Research Data

In adherence to article 29.3 of the grant agreement, beneficiaries will deposit data and associated data required to validate results presented in publications in a research data repository, within the guidelines specified in this plan.

SYNTHESYS+ has a paid agreement with the Research Ideas and Outcomes (RIO) journal, allowing publication of all project outputs including: data, methods, workflows, software, reports and articles through the life of the project. Publication in other suitable journals and/or repositories is permitted.

2. FAIR data

2.1 Making data findable, including provisions for metadata

Are the data produced and/or used in the project discoverable with metadata, identifiable and locatable by means of a standard identification mechanism, e.g. persistent and unique identifiers such as Digital Object Identifiers (DOI)?

SYNTHESYS+ will follow the persistent identification, versioning and metadata policies described in the DiSSCo DMP. This includes using the persistent URI system developed and promoted by the CETAF-ISTC group to identify specimens. For documents and datasets on GBIF and DOIs will be used. All outcomes and datasets published in RIO will be given a DOI. These will be archived and indexed in several international repositories, incl. DOAJ, OpenAire, Mendeley and others. Any specimen data made available through GBIF will also be given a DOI at dataset level, and a CETAF-ISTC persistent identifier at specimen level. For specimen data published on institutional or national portals we strongly recommend the using the persistent URI system developed and promoted by the CETAF-ISTC group.

What naming conventions do you follow?

Project reports and documentation will follow file naming conventions that reference the appropriate project work packages, tasks, milestones and deliverables. Code projects, published datasets and data architectures will follow naming conventions appropriate to the technologies and repositories used, and in line with community standards.

Will search keywords be provided that optimize possibilities for re-use?

Search keywords are included as standard on data and publications deposited in RIO. Keywords will be used in other repositories and platforms wherever possible and applicable.

Do you provide clear version numbers?

Version numbers will be used for formal reports and project documentation. Formal versioning will be used for code releases and relevant published datasets.

2.2 Making data openly accessible

Which data produced and/or used in the project will be made openly available as the default?

All project data and images will be made openly available by default under appropriate open licenses or waivers unless agreed otherwise by Project Council vote, as per the Consortium Agreement.

How will the data be made accessible?

Research data and outcomes will be made accessible through open access repositories (e.g. RIO), data portals (e.g. GBIF, NHM Data Portal) and repositories (e.g. Zenodo, GGBN, GenBank). Software source code will be deposited on GitHub. For further information see *Appendix 1*.

What methods or software tools are needed to access the data?

Any data deposited in RIO, GBIF, institutional data portals or other repositories can be accessed through a browser without the need for specialist software. Most of these also provide APIs for programmatic access to the data. Large files may be made available by arrangement through file transfer services such as B2Dropor other method that is free to the end user. An API will be developed for ELViS, however the system will also have an intuitive web-based user interface enabling data discovery via a registration and log-in system.

Is documentation about the software needed to access the data included?

Existing platforms like RIO and GBIF provide online documentation for using their software and services. For software developed within the project, such as ELViS (JRA1) and the Specimen Data Refinery (JRA3), documentation will be generated and made available to support users in accessing and using the data.

Is it possible to include the relevant software (e.g. in open source code)?

ELViS and the Specimen Data Refinery will be developed as open-source software and source code will be made available in GitHub.

Where will the data and associated metadata, documentation and code be deposited?

Research data and outcomes will be deposited, as appropriate to the type of data and content, in open access repositories (e.g. RIO), data portals (e.g. GBIF, NHM Data Portal) and repositories (e.g. Zenodo, GGBN, GenBank). Software code will be deposited on GitHub. For further information see *Appendix 1*.

Is there a need for a data access committee?

The need for a Data Access Committee has not been identified at this stage, however this decision will be reviewed in line with updates to the DMP as requirements around release of potentially sensitive collections data become clearer.

Are there well described conditions for access (i.e. a machine readable license)?

Data published on existing platforms, data portals and repositories will include a machine-readable license and other relevant access conditions where supported by the platform. Access conditions including machine readable licenses will be included in the data architecture design of ELViS.

How will the identity of the person accessing the data be ascertained?

Identifying information may be collected from individuals accessing information through ELViS (JRA1) where they have signed in with a registered user account. Data made available through VA and other datasets published under Open Data licenses or waivers will not require the individual to be identified.

2.3 Making data interoperable

Are the data produced in the project interoperable? That is, allowing data exchange and re-use between researchers, institutions, organisations, countries, etc. (i.e. adhering to standards for

SYNTHESYS+ was funded by the Horizon 2020 Framework of the European Union under the H2020 Open Innovation and Open Science Research Infrastructure call.

SYNTHESYS+ Synthesis of Systematic Resources a DiSSCo project

format; as much as possible compliant with available (open) software applications, and in particular facilitating re-combinations with different datasets from different origins?

SYNTHESYS+, along with other DiSSCo-linked projects, will follow the standards and formats of the biodiversity informatics community, and adhere to common standards and best practice among the global research and data communities. Additionally, work done under NA4 (Digital Standards & Processes) will specifically focus on linked data and interoperability. Thus best practices will be developed through the lifetime of the project and will be reflected in future iterations of the DMP.

What data and metadata vocabularies, standards or methodologies will you follow to make your data interoperable?

Common data and metadata vocabularies, standards or methodologies will be used as available and appropriate to the content and types of digital object, including ISO 19115/19139 for geographic information; ISO 1806 for dates; ISO 3166 for country codes; Darwin Core (DwC), Access to Biological Collections Data (ABCD) and its extension for geosciences (ABCDEFG), and generic standards such as Exif and IIIF for images. See section 2.1 and Appendix 1 for more details on metadata vocabularies.

Will you be using standard vocabularies for all data types present in your data set, to allow inter-disciplinary interoperability?

Standard data type vocabularies will be used and aligned with the emerging standards from the GO FAIR initiative (https://www.go-fair.org/).

In case it is unavoidable that you use uncommon or generate project-specific ontologies or vocabularies, will you provide mappings to more commonly used ontologies?

Yes, in such cases best efforts will be made to map to common shared ontologies.

2.4 Increase data re-use

How will the data be licensed to permit the widest re-use possible?

Open licenses will be used as far as possible; specific licenses will vary according to data type. Specimen data will typically be licensed under CC0, whilst images and documents will be licensed under CC-BY-4.0. Cases where specific data may warrant more restrictive licenses will be reviewed individually. For further details on data licensing please see *Appendix 1*.

When will the data be made available for re-use?

Wherever possible, non-sensitive will be made available for re-use without delay. Data embargoes are not expected, but any restrictions or exceptions arising will be discussed by the relevant project bodies (Executive Board, Project Council).

Are the data produced and/or used in the project useable by third parties, in particular after the end of the project?

Data is usable by third parties unless an exclusive license is granted according to the procedures outlined in the Consortium Agreement. Data will continue to be usable under these same terms beyond the end of the project.

How long is it intended that the data remains re-usable?

There is no time limit intended on the accessibility and re-use of the data. Data publication platforms and repositories will provide sustainability for long-term access to the data.

3. Allocation of resources

What are the costs of making data FAIR in your project?

A subscription to the Research Outcomes and Ideas journal (RIO) is agreed for the duration of the project, at a one-off cost of 6,665 EUR. These costs will be covered by the Horizon 2020 grant, under the NA1 (Management) budget. Other costs of making data available may be supported through the VA1 budget, other personnel time, or in-kind contributions by participating institutions.

Who will be responsible for data management in your project?

Data management falls under the Management (NA1) work package and is the overall responsibility of the project co-ordinator (NHM). However, participating institutions will be responsible for the management of the data they individually generate, ensuring this is consistent with the SYNTHESYS+ and DiSSCo DMPs.

Are the resources for long-term preservation discussed (costs and potential value, who decides and how what data will be kept for how long?)

Sustainability of much of the data preservation will be assured by developing ELViS as open source software and transferring maintenance to the DiSSCo consortium after the project, providing a future development path and adoption by the wider network of DiSSCo partners. In general, the DiSSCo roadmap will seek to ensure long-term preservation of data from all related projects.

4. Data Security

Security provisions will be implemented at the institutional level. Transnational Access data held at the NHM is subject to security precautions outlined in D11.1. Security provisions applied to the European Loans and Visits System will be further detailed in future versions of this DMP.

5. Ethical aspects

SYNTHESYS+, specifically WP9 (TA1) involves the processing of personal data. Ethical considerations in relation to personal data processing are covered in Deliverable D11.1.

6. Other issues

Do you make use of other national/funder/sectorial/departmental procedures for data management? If yes, which ones?

All partners may be subject to institutional data management procedures and/or national regulations in addition to this Data Management Plan.

SYNTHESYS+ was funded by the Horizon 2020 Framework of the European Union under the H2020 Open Innovation and Open Science Research Infrastructure call.

Appendix 1: Data Types Summary Table

Data type	Category	Store	Publication	Personal data	NA1	NA2	NA3	NA4	NA5	JRA1	JRA2	JRA3	TA1	VA1	Licenses*
Project reports, specifications and requirement docs	Documents	File stores	RIO		created	created	created	created	created	created	created	created	created	created	CC-BY-4.0
Papers	Documents	File stores	RIO/Other Open Access		created	created	created	created	created	created	created	created	created	created	CC-BY-4.0
Policy documents	Documents	Filestores	RIO/Other Open Access	yes		held							created		CC-BY-4.0
Surveys	Datasets	Databases	Various												N/A
Training and policy material	Documents, web pages	Filestores, web platforms	Various			created	created								CC-BY-4.0
Specimen data	Datasets	Databases	Data Portals				held			held		held		created	CC0
Specimen images	Files	File stores	Data Portals							held		held		created	Varied (mostly CC-BY)
3D specimen images and datasets	Files, datasets	File stores	Data Portals								created			created	Varied (mostly CC-BY)
Specimen analysis data	Datasets	Databases	Data Portals								created	created			CC0
Molecular data	Datasets	Databases and/or filestores	GGBN Portal (Unless restrictions apply)								created				CC0
Loan, visit and ticketing data	Datasets	Databases	N/A	yes						created			created	created	N/A
Code and models	Files	Repositories	GitHub									created			GPL-2.0 or similar
Institution data	Datasets	Databases	ELViS			held	held			held			held	held	CC0
Collections data	Datasets	Databases	ELViS			created	created			created			held	held	CC0
Facility data	Datasets	Databases	ELViS				created			held			held	held	CC0
Staff and/or user data	Datasets	Databases	N/A	yes						created			created	created	N/A
User stories	Documents	File stores	GitHub							created					GPL-2.0 or similar

Table 1: Summary of SYNTHESYS+ Data Types and Associated Information

*Expected licenses – exceptions may apply in all categories

[See here to access full table](#)
(Google Sheets)

SYNTHESYS+ was funded by the Horizon 2020 Framework of the European Union under the H2020 Open Innovation and Open Science Research Infrastructure call.