DiSSCo related output

This template collects the required metadata to reference the official Deliverables and Milestones of DiSSCo-related projects. More information on the mandatory and conditionally mandatory fields can be found in the supporting document 'Metadata for DiSSCo Knowledge base' that is shared among work package leads, and in Teamwork > Files. A short explanatory text is given for all metadata fields, thus allowing easy entry of the required information. If there are any questions, please contact us at info@dissco.eu.

Title
Deliverable 9.3 1st DiSSCo Prepare All Hands Meeting

Author(s)
Eva Alonso

Dimitris Koureas

Identifier of the author(s)
https://orcid.org/0000-0001-5336-9723

https://orcid.org/0000-0002-4842-6487

Affiliation
Naturalis Biodiversity Center

Contributors
Dimitris Koureas, Ana Casino, Wouter Addink, Marie-Laure Kamatali, Sharif Islam, Neil Raes, Tina Loo, Henrik Enghoff, Mareike Petersen, Aino Juslén, Sabine von Mering, Rui Figueira, Elsa Fontainha, Helen Hardy, Vince Smith, Laurence Livermore, Matt Woodburn, Carole Paleco, Serge Scory, Alex Hardisty, Claus Weiland, Falko Glöckler, Ohad Graber-Soudry, Patricia Mergen, Michel Guiraud, Salomé Landel, François Dusolier, Jose Alonso, Julia Pim Reis, Eva Pérez, Judite Alves.

Publisher

Identifier of the publisher

Resource ID

Publication year
2021

Related identifiers

Is it the first time you submit this outcome?
Yes

Creation date
29/01/2021

Version
1

Citation
DiSSCo Prepare -AHM1

Abstract
DiSSCo's first all-hands meeting (AHM1) took place from 18-22 January 2021 as an online event hosted by the Naturalis Biodiversity Center as coordinator of DiSSCo Prepare.
The objective of the meeting was two-fold. First, to collect and share information about the project's status, through a series of work-intensive sessions. Second, to build on the overall understanding necessary to foster further discussion on how to optimise DiSSCo Prepare's performance. The discussions took into

account the multiple synergies but also dependencies and risks present in the work program.

Content keywords
other

Other content keywords
management, AHM

Project reference
DiSSCo Prepare (GA-871043)

WP number
WP9

Project output
Deliverable

Deliverable/milestone number
9.3

Dissemination level
Public

Rights

License
CC0 1.0 Universal (CC0 1.0)

Resource type
Text

Format
pdf

Funding Programme
H2020-INFRADEV-2019-2

Contact email
eva.alonso@naturalis.nl

DISSCO PREPARE AHM1
Deliverable 9.3

January 18-22, 2021

Summary

1. Introduction

DiSSCo's first all-hands meeting (AHM1) took place from 18-22 January 2021 as an online event hosted by the Naturalis Biodiversity Center as coordinator of DiSSCo Prepare.

The objective of the meeting was two-fold. First, to collect and share information about the project's status, through a series of work-intensive sessions. Second, to build on the overall understanding necessary to foster further discussion on how to optimise DiSSCo Prepare's performance. The discussions took into account the multiple synergies but also dependencies and risks present in the work program.

The AHM1 comprised 21 consecutive sessions over five days, including the inaugural meeting of DiSSCo Prepare's Project Council that also took place during the event. Despite being a virtual event because of the pandemic, the sessions attracted a high number of attendees who had the opportunity to thoroughly discuss all of DiSSCo Prepare's work packages.

To motivate the participants to get actively involved in the discussions, the opening session on day one started with a keynote speech by Dr. Isabel Sousa Pinto (IPBES), who talked about the crucial contribution of European RI(s) to achieve objectives of global biodiversity.

Key facts of the event

1. Deliverable 9.3. Led by Naturalis. Due date, January 2021.
2. The virtual meeting took place from 18 to 22 January 2021. It comprised 21 consecutive sessions.
3. Wide international participation: A total of 120 participants registered and attended the sessions, fully or partially. These included members from all 30 beneficiary partners of DiSSCo and external experts from other institutions (i.e, CNRI).
4. Sessions registered an average of 40 participants per session.
5. 51% of attendees to the plenary session were women.
6. The AHM1 hosted the inaugural meeting of the DiSSCo Prepare governing body, its Project Council.
7. The results of a post-event survey showed general satisfaction with the event. The attendees perceived it as an excellent way to move forward.

The AHM1 in a nutshell

The AHM1 consisted of 20 consecutive 90-minutes working sessions -including a plenary session on the first day and a wrap-up session at the end of the event- that allowed participants to get better knowledge and contribute meaningfully to topics that may fall out of their responsibilities in the project.

This chapter summarises the key outcomes from the working sessions, providing a view of the remarkable progress resulted from the discussions.

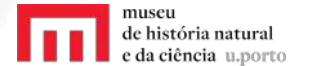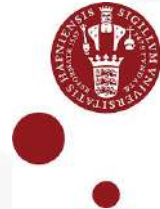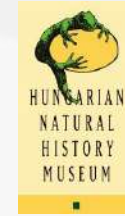| Agenda		
9:00-9:10	**1. Quorum & Agenda**	DK
9:10-9:55	**2. DPP update on progress**	DK/Stream Coordinators
9:55-10:05	**3. Risk registry update**	EA
10:05-10:25	**4. Extension of Partner's participation**	DK
10:25-10:30	**5. AoB**	All

1 February 2020 – 31 January 2023

21 Countries
(DiSSCo Countries)

30 Beneficiaries
198 Individuals

EUR 5 million
(approx. 20% in-kind contributions)

Implementation Timeline / Funding structure / Governance

How we work through the Preparatory Phase of DiSSCo
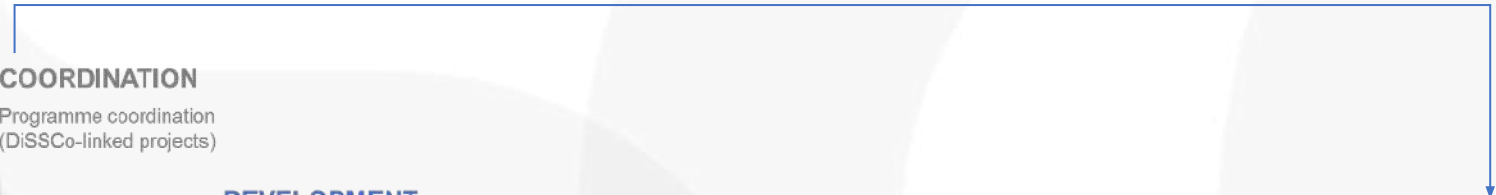
Our goals

Improve the DiSSCo **Implementation Readiness Level (IRL)** and

Deliver DiSSCo **Construction Masterplan**
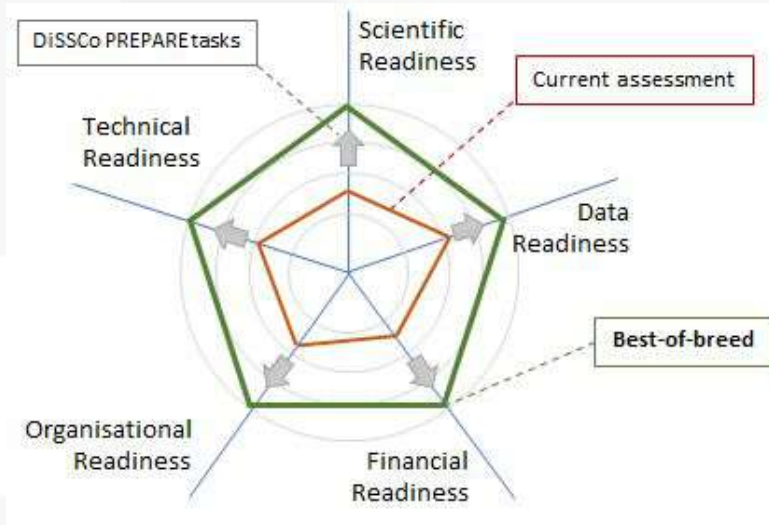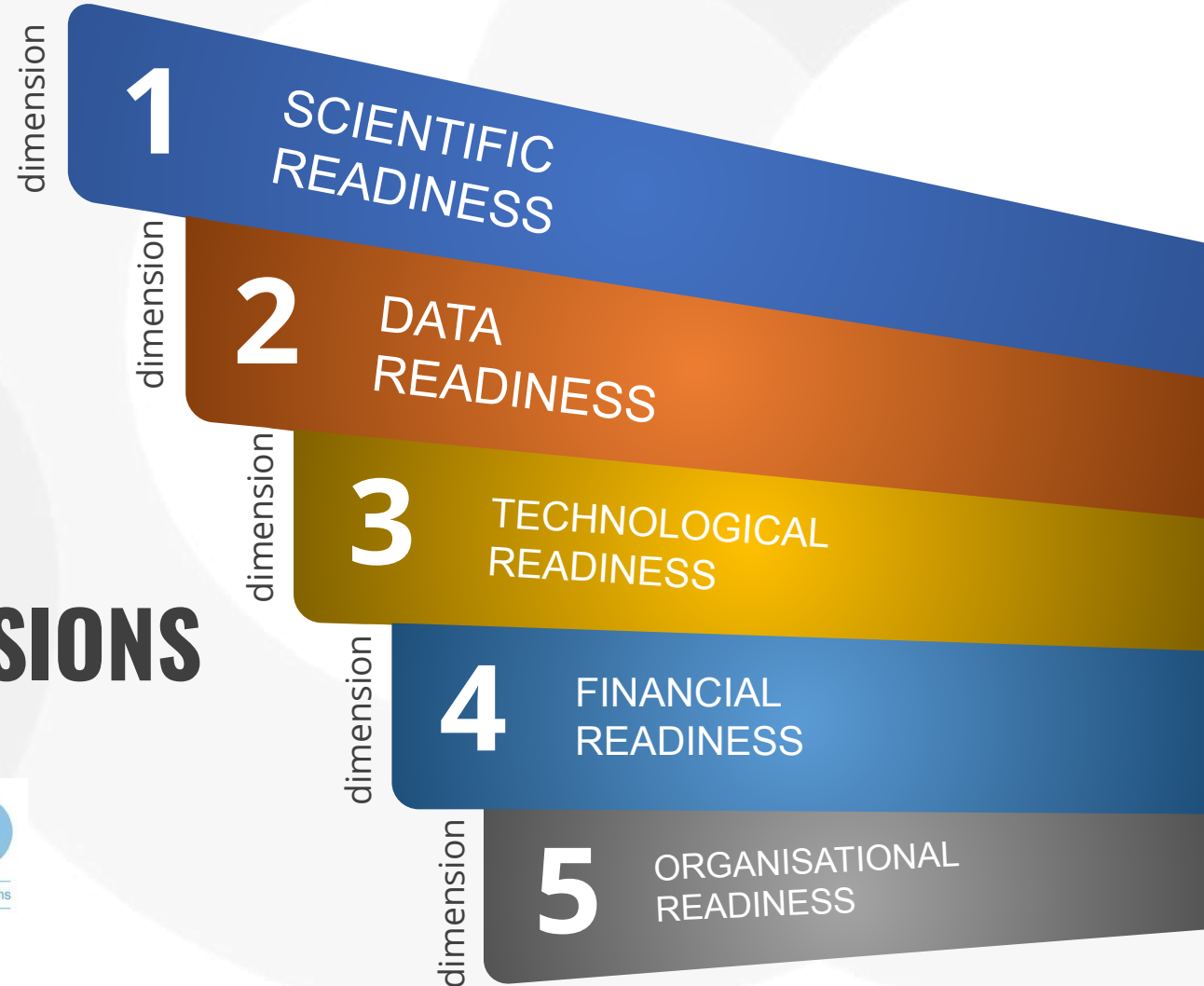
36 Deliverables
70 Milestones

DIMENSION

1

1 SCIENTIFIC READINESS

- Construct a **service development framework** focused on users in NSC-related research and research applications;

- Identify the criteria for establishing **priority for the digitisation**, data generation and enrichment of NSCs;

- Put together a **socioeconomic benefits framework** for partners and countries;

- Produce a **training strategy** to address identified needs;

- Design a **helpdesk** that will provide DiSSCo with the necessary user support services;

- Provide DiSSCo with a **human resources policy**.

DIMENSION
2

- Describe the mechanisms and tools to **improve digital skills and competencies** across facilities;
- Collate, refine and implement **best practices for data generation, enrichment and mobilisation at the institutional level**;
- Develop **secondment and distributed team working practices**.

2 DATA READINESS

DIMENSION

3

- Build a **knowledgebase with structured and validated tools and products** relevant to the DiSSCo RI;
- Provide a **modelling framework and data model** covering all requirements from the natural science collections domain in alignment with the broader biodiversity research data domain;
- Provide guidelines for (machine to machine) **data enhancement** allowing for cross-linking of information and increased interoperability (FAIR);
- Provide construction **plans for key services** for seamless integration into the overall technical DiSSCo Architecture;
- Refine the **DiSSCo technical architecture for digital specimen data** and provide an overall plan for implementation and deployment of this architecture;
- Integrate **DiSSCo in the technical landscape** of national, European and global biodiversity data providers.

3 TECHNOLOGICAL READINESS

DIMENSION

4

- Provide DiSSCo with a **sound knowledge of its cost structure** in order to estimate running costs and costs to be charged to users for digitisation-on-demand;

- Draw up the **business model for DiSSCo at a national and international level** based on governmental and institutional contributions, users' charges and industrial sector contribution through R&D projects.

4 FINANCIAL READINESS

DIMENSION

5

- Provide DiSSCo with a **ready–to–implement organisational model** at a European and national level;
- Provide DiSSCo with a clear and efficient pathway towards its **establishment as a legal entity**;
- Provide DiSSCo with a **common set of access and management policies**;
- **Liaise with national governments** and external stakeholders to ensure their commitment, gather their input to enhance DiSSCo, and **establish partnership frameworks**;
- Develop recommendations to distribute the **specialisation towards collections digitisation, generation and enrichment**;
- Operate **effective communication and dissemination tools** to raise awareness and ensure the uptake of project results.

5 ORGANISATIONAL READINESS

DiSSCo Prepare project governance
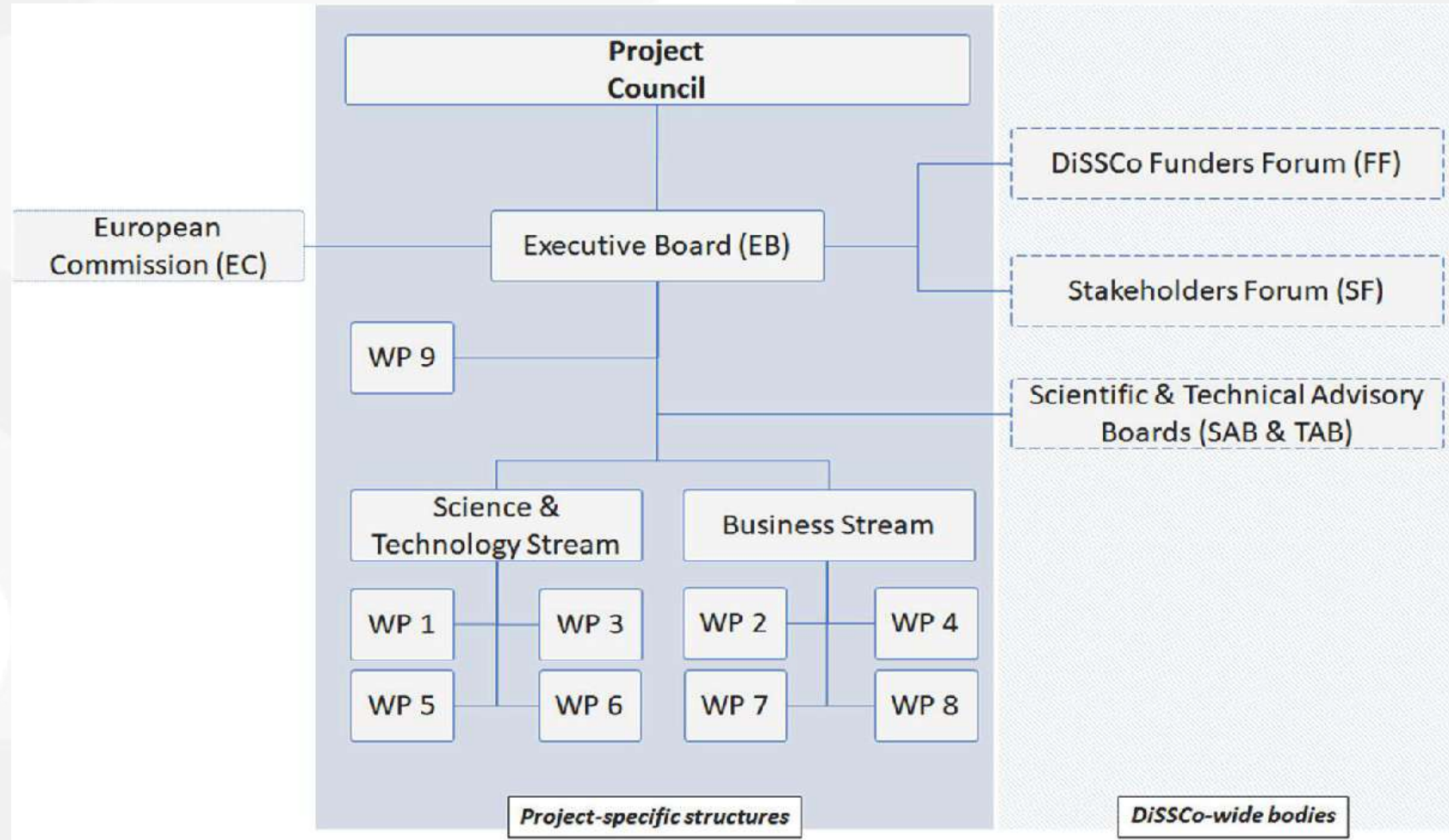
DiSSCo Prepare dependencies to other DiSSCo-linked projects
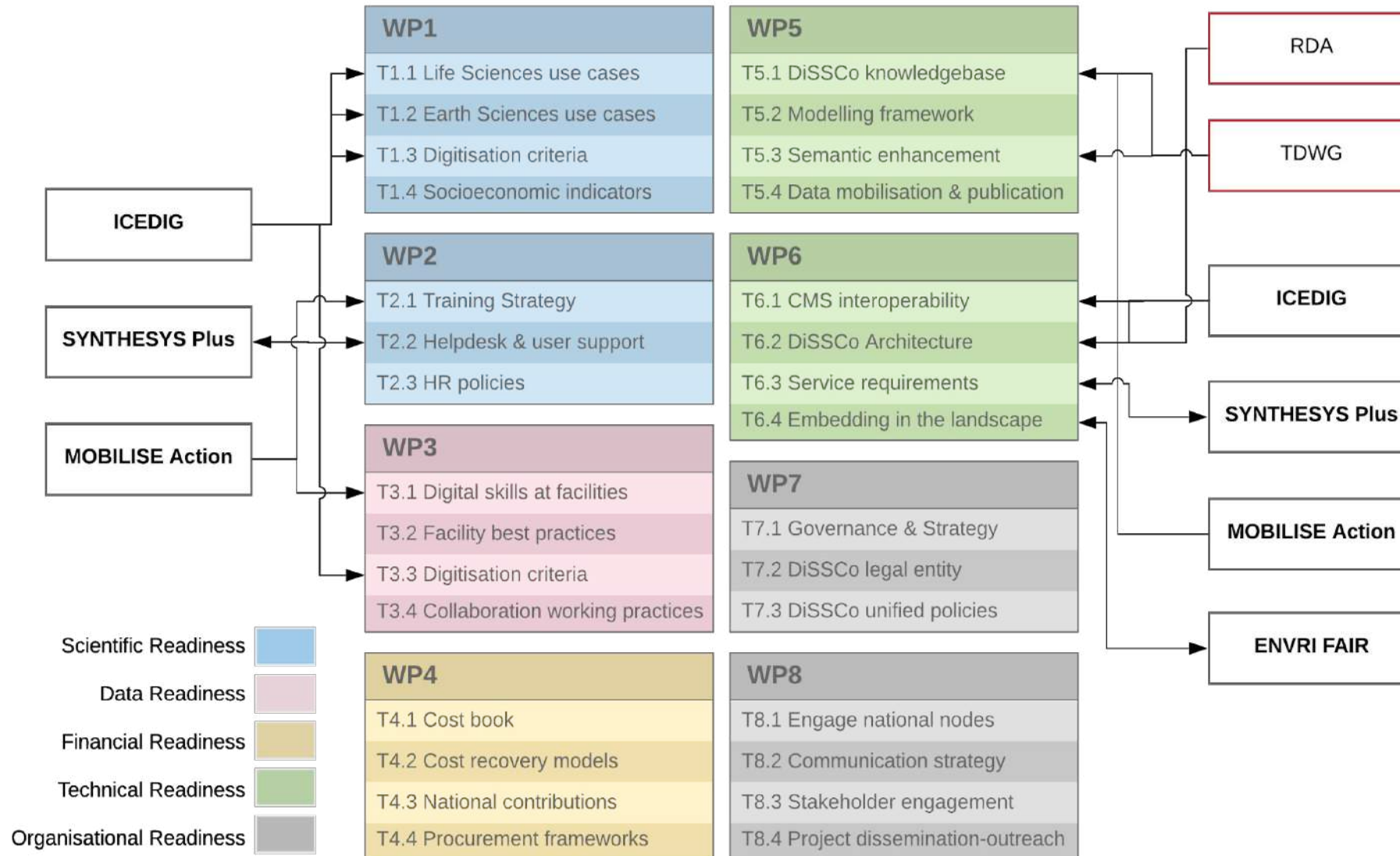
First outputs
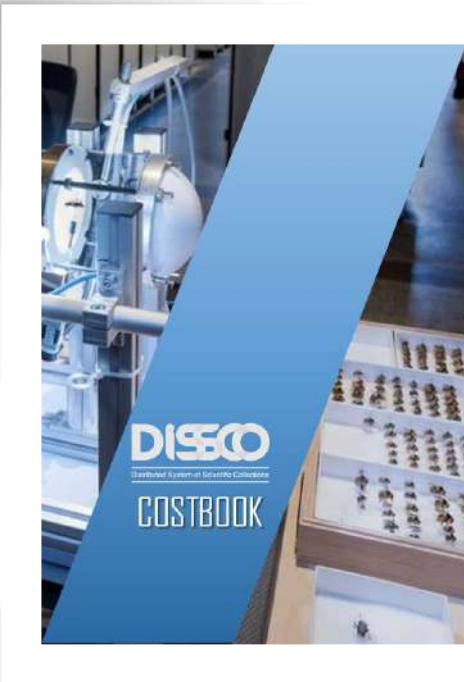
Main Cost
Parameters identified

WP4

OpenDS and MIDS
standard development

WP6

Communication
Strategy

WP8

LE options set
and now evaluated

WP7

DPP Risk registry

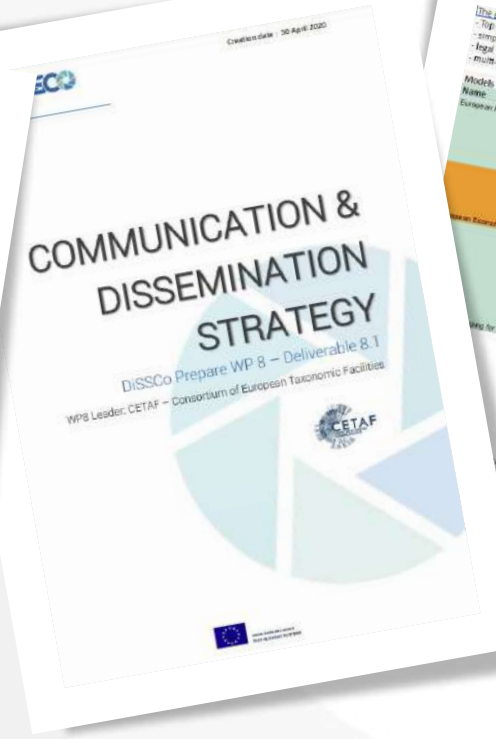| #	Source	Created By	Probability Low	Probability Medium	Probability High	Impact Low	Impact Medium	Impact High	Result	Impact Areas Cost	Impact Areas Schedule	Impact Areas Performance	Mitigation/Response Plan	Status
1	Unable to engage with industrial stakeholders. The community of DiSSCo lacks strong ties with industry, and the specificity of the DiSSCo content may drive to a lack of attractiveness for industrial partners to become engaged in any partnership. WP8	Eva A. 18/03/2020			8			8	64			x	To entice them to participate, it will be necessary 1) to identify a clear realm of potential areas where industrial partners may have an interest for further innovation and developing a commercially feasible final produce, and 2) to prepare a set of business cases where the partnership museum-industry has proven efficient and successful.	Open
2	No alignment between Work packages. WP4 relies on results from other work packages in order to meet its goals. Delays in designing services will affect the design of the Cost Book.	Eva A. 18/03/2020		5				8	40			x	WP4 is to be involved in the initial stages of design services.	Open
3	Difficulty to harmonising positions throughout the NTFs. The DiSSCo partnership includes a high number of countries that strongly differ in terms of political and financial commitment as well as regarding their internal development in the matter (level of integration of collections in national RI roadmaps, the degree of maturity of the teams involved, etc.) WP8	Eva A. 18/03/2020			8		5		40			x	Strong, harmonised efforts will be devoted from the very beginning by deploying specific actions to upgrade the less-developed partners and to having the more-developed ones to practice a continuous benchmarking and to provide permanent support to the others (MOBILISE COST Action and SYNTHESYS+ will be essential for those purposes).	Open
4	Lack of taking ownership of DiSSCo at national level. For the success of DiSSCo, commitment in member countries is of utmost importance. This refers to the different governmental levels as well as to the institutional one.	Eva A. 18/03/2020		5				8	40			x	DiSSCo Prepare will provide the resources and the necessary guidance to the national nodes as to ensure channelling the necessary information to the involved governmental parties (to make sure the clear messages to through), maintaining a permanent close contact with national representatives (to meet their expectations and concerns) and ensuring that all parties take ownership of DiSSCo.	Open
5	Dependencies to external projects. Some tasks in the project build on top of outputs of other external projects. Low quality or failing to deliver in those projects might affect the quality of the DiSSCo Prepare outputs.	Eva A. 18/03/2020		5				8	40			x	Through the Strategic Alignment Group (coordinators of all DiSSCo-linked projects) the DiSSCo coordination office will ensure a birds-eye view to all the work performed across projects. This allows for a programmatic view of all projects. Risk mitigation procedure at the programme-level is put in place to mitigate issues and ensure complementarity between similar tasks across projects, including DiSSCo Prepare.	Open
6	Insufficient material for analysis. All WP1 activities primarily concern analysis and synthesis of previous studies. A general risk is that one or more of the complied corpora of studies proves insufficient for meaningful analysis.	Eva A. 18/03/2020		5			5		25			x	Acquisition of additional material, through targeted questionnaire surveys and/or in-depth interviews of key persons.	Open
7	Data source availability. Lack of good data sources for automated institutional capacity/skill harvesting. T3.1	Eva A. 18/03/2020		5			5		25			x	Manage a list of data sources and invest in community efforts to explain the need to publish these data in a reusable form.	Open
8	Governance of software adoption processes. No effective governance structures for discussing and implementing the adoption of common software platforms affecting the WP3 deliverables.	Eva A. 18/03/2020	1					8	8			x	Develop the necessary agreements among WP3 partners.	Open
9	Distributed development team pilot. Timelines for the pilot of distributed development teams do not align with WP6/ DiSSCo RI development activities.	Eva A. 18/03/2020	1					8	8		x	x	Work closely with WP6 and be flexible with the timing of the pilot delivery.	Open
10	Timelines of internal developments do not align with WP6/DiSSCo RI. The core task of WP6 -design and planning of the DiSSCo technical architecture is directly dependent on the preparatory work in WP5 on the data model, standards, etc. Expected is a continuation and extension of the prior work in ICEDIG.	Eva A. 18/03/2020	1					8	8		x	x	Key persons in WP6 are also actively contributing partners in WP5 (and vice versa) to ensure linkage and synchornisation in both technical WPs.	Open
11	Unresponsive partner. A partner becomes unresponsive during the project, jeopardising the delivery of outputs. WP9	Eva A. 18/03/2020	1					8	8		x	x	The partner is enquired by the Project Manager. The issue will be brought to the Executive Board (EB) for action if still unresolved. The partner is replaced by another Consortium member in the same WP when decided by decision of the Project Council.	Open
12	"No legal entity model fits all" Some participating institutions or countries can not join the proposed legal entity. WP7	Eva A. 18/03/2020	1				6		6			x	Detailed analysis of all implications of the various models to form a legal entity, and frequent dialogue with all the stakeholders.	Open
13	Delayed project output affects project dependencies. A delayed task output might cascade to other tasks because of internal project dependencies. WP9	Eva A. 18/03/2020		5		1			5		x	x	Dependencies in the project are not end-start dependencies, rather end-end dependencies. This means that delayed (for less than 6 months) outputs should not affect the delivery of the work in the dependent tasks.	Open

ALL HAND MEETING - AHM1
Business Stream

Ana Casino - CETAF

DiSSCo Prepare

Business Stream Coordinator

Collaborative work
Business Stream

Intense, challenging, effective, successful

Meetings: 12 BS meetings, numerous others

Milestones: 5 MSs achieved +3 in short, 2 delayed

Deliverables: 1 D produced

BUSINESS STREAM
Timeline
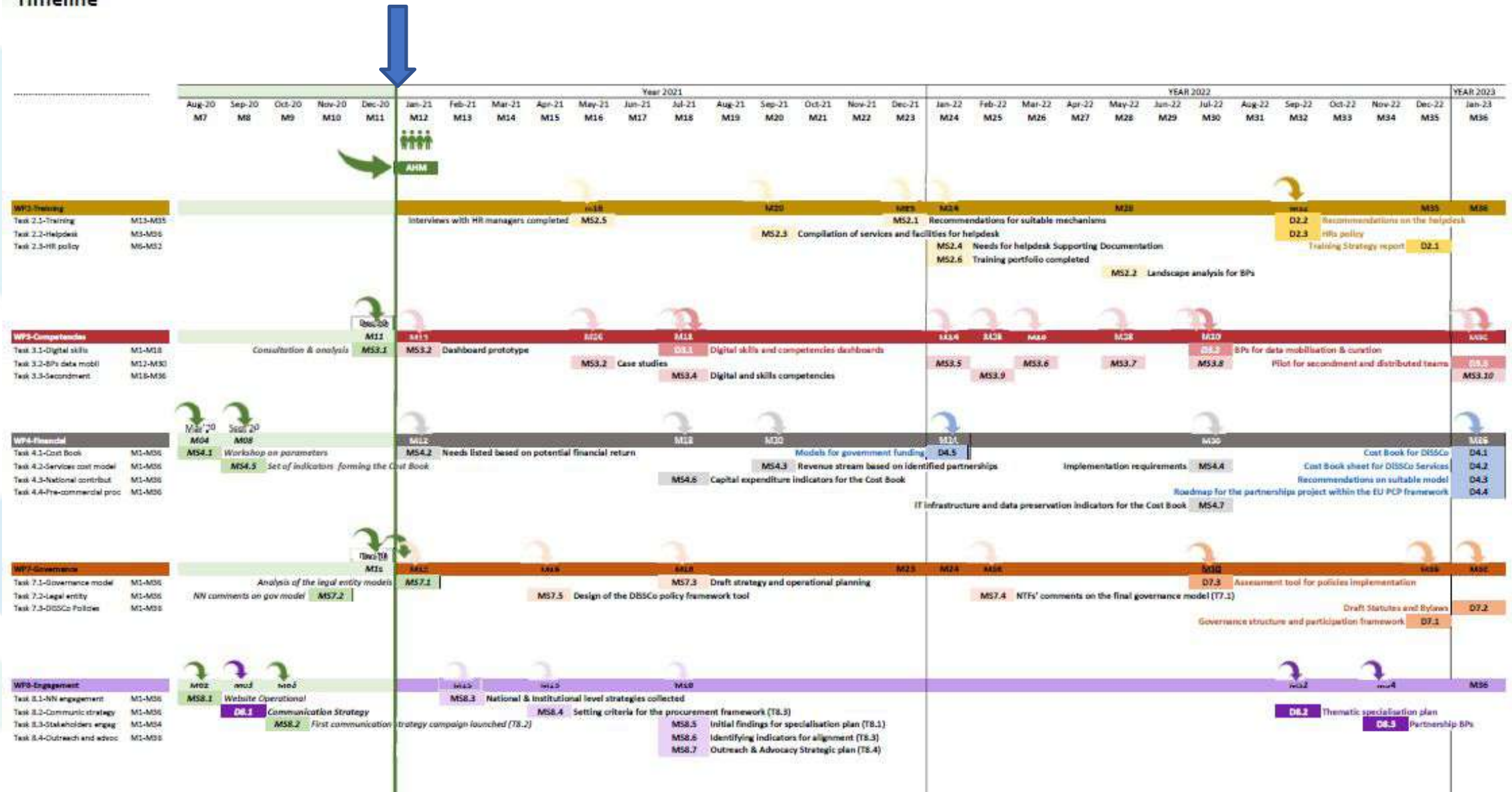
Progress made in WP2 (ULISBOA)

Human Resources, Training & Users Support

- T2.1-Training Strategy, just to start (M12), KoM-22 Jan 21
- T2.2-Helpdesk, needs to leverage on SYNTHESYS+ and DPP WPs
- T2.3-HRs policy, start has been delayed (from M6 to M11)

Achievements in 2020: no MS or D expected

- Defining **scope, participants and contributions**
- **Delay** for MS2.5 (T2.3) and a change of scope
- Establishing **working flows** of information
- Strong **coordination** needed with SYNTHESYS+ T2.3, T2.4 and DPP WP3,5

T2.2 Helpdesk Support to users

offering information and supporting users that will use DiSSCo services and facilities

D2.2 Recommendations on Helpdesk design (Sept'22)

- compilation of services and facilities to be offered

- identification of different modalities of interaction with users and stakeholders

- identification of supporting documents needed

- required features (RIs, networks and organizations to interact)

Timeline:

MS2.3 "Compilation of services and facilities" (Sept'21)

MS2.4 "Supporting documentation identified" (Jan'22)

Challenges:

Connection to outcomes from ICEDIG, SYNTHESYS+ and DPP WPs 3, 6, 7

T2.3 HRs policy definition
RI-related human resources have the appropriate skills and abilities

D2.3 DiSSCo HR policy (Sept'22)

- new task leader on board

- rescheduling of work and redefinition of sub-tasks:
    * identification of factors that induce strategic performance in HRM
    * landscape analysis of policies at ESFRI /ERIC level
    * portofolio of roles and competencies necessary to support DISSCo operation
    * Definition of HR practices (training, performance and team-based organisation)

Timeline:

MS2.5 "Interviews with HRs departments at similar RIs" (May'21)

MS2.6 "Portfolio of competencies and roles" (Jan'22)

Challenges:

- How to best integrate results from other projects

- How to harmonise positions throughout NNs regarding traiing needs

- Strong technical component (WP6-Techncial architecture and services)

WP2 – Training, HR policies and helpdesk support

Take-home messages
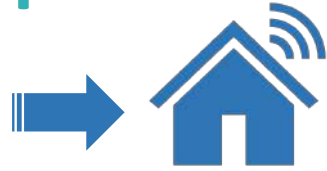
- Tasks are **running** according to plan

- Upcoming 6 months are crucial and some speed up is needed
  - Launch and get on track for T2.1 in alignment with SYNTHESYS+ **Training Roadmap**
  - Conclude the definition of **services** to be considered in the Helpdesk (T2.2).
  - Start elaborating on the **needs of supporting documentation** to be included in the Helpdesk (T2.2)
  - Perform **interview**s with relevant Ris and conclude the ERIC landscape analysis on HRs (T2.3)

- **Communication** with WPs (WP3, WP6, WP7, but also WP1) should continue and be strengthened
  - attending meetings
  - reviewing reports

- Same **coordination** is required for Synthesis+ (Tasks NA2.3 and NA2.4)

Progress made in WP4 (MNHN)

Business framework

All 4 Tasks at full speed

Achievements in 2020: **2 MSs** achieved /no D expected

MS4.1 Workshop on Cost Book parameters (May'20)
MS4.5 Set of indicators forming the basis of Cost Book (Sept'20)

- Definition **Cost Book structure**
- **Questionnaire** sent to NNs
- Development of **Cost methodology**
- **Procurement systems** – case studies
- Intense **communication:** Bilateral meetings (11), interviews (38), recurrent consultancy

T4.1 Cost Book for DiSSCo
Thursday, 21 January (9h00-10h30) – Michel Guiraud (MNHN)

D4.1 DiSSCo Cost Book (Jan'23)

Methodology assessment for indirect costs

- Identification of Indirect Costs
- Base references: sqm and personnel
- Proposal of alternative methods:
  - a) In-house calculations
  - b) Flat ratio (equal across partners)
  - c) Flat ratio (weighted by type or other criteria)

Timeline:

MS4.6 "Capital expenditure indicators for the Cost Book" (July'21)

Challenges:

- Combination of **"simplicity" and "robustness"** requirements
- Alongside **variety of typology and accounting** systems
- Lack of communication lines between **scientific and administrative/financial** departments

T4.2 Cost model for charging services
Thursday, 21 January (13h00-14h30) - Laurence Livermore (NHM)

D4.2 Cost Book sheet for DiSSCo Services (Jan'23)

Scoping exercise for Cost Book sheet

- Definition of scope, components and dependencies (T1.4)

- Identification of sub-tasks:

    4.2.1 Digitization costs (including mobilization, data processing and data access and exploitation)

    4.2.2 Common services to consider: Helpdesk, training, communication, HR and others

    4.2.3 Cost model platform

    4.2.4 User's needs and testing

- Identification of subcontractor (with BGM and MNHN)

Timeline:

MS4.2 Needs listed based on potential financial return" (Jan'21)

Challenges:

- Strong technical component (WP6-Techncial architecture and services)

- Alignment with web-based tool for policies-related assessment (T7.3)

WP4 – Financial framework

Take-home messages

- WP4 aims at overcoming the **diversity of countries and institutions**

- It will require exchanges between **scientific and administration** people, which constitutes a huge challenge

- Member's representatives will be requested
    - provide information **beyond their usual competencies**
    - identify **relevant people**

- The DiSSCo Cost Book will be kept as simple and robust as possible!

Progress made in WP7 (RBINS)

Governance, Policy & Legal frameworks

- T7.1-Governance model is being structured
- T7.2-Legal entity definition work is at full speed
- T7.3-Policies framework is to start

Achievements in 2020:

MS7.2 - Summary of NNs comments on the amended EU MoU (Oct'20)
MS7.1 - Recommendations on DiSSCo legal entity model (Dec'20)

- Analysis of **criteria and parameters** of the model to best suit DiSSCo
- In-depth **reporting** on existing models and their **use and feasibility**
- Regular **advisory meetings** with external legal expert
- Identification of the **legal entity** proposed for approval by the governing bodies

T7.1 Governance model

Monday, 18 January (15h00-16h30) – Dimitris Koureas, Eva Alonso (Naturalis)

D7.1 Plan for governance structure and function (Feb'22)

DiSSCo Governance model-requirements and landscape analysis

- Definition of the **scope**:
  * A governance chart and ToRs for the various bodies
  * A mid-and longer-term strategy for DiSSCo & an operational planning
- Identification of **requirements**
- Preparation of the Landscape analysis and the methodology to follow, from criteria to met to BPs from similar Ris

Timeline:

MS7.3 "Draft strategy document and operational planning framework" (advanced to Jun'21)

MS7.4 "Summary of comments on the updated governance model" (advanced to Nov'21)

Challenges:

- Strong interdependencies between LE and GM

T7.2 Legal entity model
Wednesday, 20 January (15h00-16h30) – Serge Scory, Carole Paleco (RBINS)

D7.2 Draft statutes and by-laws,implementation plan (Dec'21)

DiSSCo Legal entity

- **Proposed** (AISBL and ERIC) and finally **recommended** model (ERIC)
- Consideration to integrate:
  - * Role of the **community/CETAF** differentiated from others (as GBIF)
- **Lessons learnt** from EOSC and experience gained from ERIC Forum (project)
- Alignment with **external review** (FF) and **approval** by iGA3

Challenges:
- Strong interdependencies between LE and GM
- Governance structure to allocate all needs (incl. community/CETAF participation)
- Feedback to collect from FF which may influence the recommended model (national constraints)

T7.3 Policy assessment tool
Wednesday, 20 January (13h00-14h30) – Vince Smith, Matt Woodburn (NHM)

D7.3 Assessment tool for policies implementation (July'22)
Policy Framework (Self-assessment) Tool

- Definition of **scope, requirements, terms and classification** (SYNTHESYS+)
- Leverage on **metada schema** for DiSSCo services (ELViS)
- Identification of **user stories** (29) to follow services provision
- Definition of **workflows** towards required interfaces (APIs)

Timeline:

MS7.5 "Design DiSSCo policy framework" (Apr'21)

Challenges:
- Incidence of Brexit
- Variety of areas, disparity of frameworks for policies implementation
- Surplus of web-based tools (sustainability, maintenance, update)

WP7 – Legal and organizational framework

Take-home messages

- **Advances** have been made
  - approval of a new **governance model** for the preparatory phase (iGA);
  - proposal of a **legal entity model**, 'fit for purpose', after a thorough analysis;
  - definition of a **tool to collate** relevant policies and to enable compliancy assessments

- **Partners, NNs and Stakeholders** are urged to have a critical but constructive review

- **Next steps** for the upcoming months:
  - defining the statutes and by-laws of the legal entity;
  - defining the corresponding governance model;
  - feeding the "policy assessment tool" with relevant information.

- **Business stream is essential** to gain a common understanding on the objectives of the overall project and get inspired

Progress made in WP8 (CETAF)

Stakeholder engagement & Communication Strategy

- T8.1-Engagement, is on track and proves strong support
- T8.2- Communication Strategy, ongoing process
- T8.3-Stakeholders' engagement, has not started yet
- T8.4- Advocacy strategy, owards the FF

Achievements in 2020:

MS8.1 – Website operational (March'20)

MS8.2 –First communication campaign launched (Sept'20)

D8.1 – Communication and Dissemination Strategy (Apr'20)

- Communication and Advocacy strategies in place
- Development of communication tools (website, brochure, key messages)
- Engagement with NNs, national strategies and priorities, fact sheets, bilateral meetings
- Setting the FF / Inaugural meeting

WP8 Stakeholders engagement and comm strategy

Thursday, 21 January (15h00-16h30) – Ana Casino (CETAF)

D8.2 Thematic Specialisation Plan(July'22)

- Shared vision on the basic statement
- Definition of domains, areas of influence
- Identification of a workplan to collate information, analyse and classify it, produce the tool

Timeline:

MS8.3 "collection of institutional strategies and policies" (Feb'21)
MS8.5 "Initial findings for the Specialisation Plan" (Jul'21)

Challenges:

- Taylor-made solutions from DiSSCo overall vision to national priorities/requirements
- For the Specialisation Plan:
    * Overcomplexity
    * Long-term approach
    * Parallel/complementary endeavor (with e.g. T8.3 Policies framework)

WP8 – Communication, engagement and advocacy

Take-home messages

- Need to **c**ontinue (and increase) supporting NNs, despite current cosntraints

- **Task T8.3** needs to get at speed and leverage on ongoing work (NA5 in SYNTHESYS+)

- Communication tools to serve **disseminating a clear vision** of what DiSSCo is and will be for both internal partners and external stakeholders

- **Next steps** for the upcoming months:
  - Launching the Specialisation roadmap (domain, criteria, requirements);
  - Thorough revision of the DiSSCo website and implementation of material developed;
  - Funders Forum inaugural meeting (Feb'21);
  - Support to the iGA3 (June, in Florence, IT)

- Constructive discussion needs to drive **inclusion of the community** in the governance model: Integral, inclusive, innovative

Progress made in WP3 (NHM)

Capacity Enhancement

Task T3.1 Improve digital skills and competencies across DiSSCo facilities

Achievements in 2020:

MS3.1 Improving digital capability. Case studies and analysis (Dec'20)

Participation in Stream Coordination meetings, feedback compiled, close linkage to other developments (WP2)

Overall shared insights
Business Stream

Participation of NNs is instrumental, feedback is much needed

As a distributed RI, DiSSCo needs to root on national priorities, to involve institutional requirements, but equally channel scientific community needs

Quality (and added value) of expected outcomes may imply to facilitate strong interconnectivity, align interdependencies, refine task scope and adjustment of timelines

COVID-19 pandemic has put a lot of stress on partners' staff but, on the other hand, has a beneficial multiplier effect (attendance to other WPs, insights share, acknowledgement of others' results, challenges and solutions,…)

Thanks for your attention!

Any question?

"Bringing the irreplaceable data stored in natural science collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

The preparatory phase project of DiSSCo Research Infrastructure - Distributed System of Scientific Collections

https://www.dissco.eu/prepare/

Project Council meeting 1
Update on progress - Science and Technology

Wouter Addink- Naturalis

S&T Stream leader

21 Jan 2021

Science & Technology Stream

Workpackages: 1, 3, 5, 6

Stream objectives

The WPs in this stream aim to improve the DiSSCo's Implementation Readiness Level for:
Scientific Readiness, Technological Readiness and Data Readiness

- **Scientific Readiness:** Capacity of the RI to respond/ adjust to current and anticipated user needs
- **Technological Readiness:** Capacity of the RI to meet the functional requirements of its users through comprehensive and sustainable technological solutions
- **Data Readiness:** Capacity of the RI data produces and stewards to serve FAIR and enriched data

The Work Packages

WP1: User needs and Socioeconomic impact

Related project goals (Scientific Readiness):
- Construct a service development framework focused on users in NSC related research
- Identify criteria for prioritization in digitization, data generation and enrichment
- Develop a socioeconomic benefits framework for partners and countries

- Work focus **M1-12: compiling life science and earth science user stories and use cases**

WP3: Capacity Enhancement

Related project goals (Data Readiness):
- Describe mechanisms and tools to improve digital skills and competencies
- Collate and implement best practices for data generation, enrichment and mobilization
- Develop secondment and distributed team working practices

- Work focus **M1-12: T3.1 towards D3.1 Digital skills and competencies dashboards**

The Work Packages (2)

WP5: Common resources and standards
WP6: Technical architecture & service provision

Related project goals (Technological Readiness):
- Build a knowledgebase with structured and validated tools and products relevant to DiSSCo
- Provide a modelling framework and datamodel covering all requirements from the NSC domain
- Provide guidelines for (machine to machine) data enhancement allowing from cross-linking and interoperability
- Provide construction plans for key services for integration with the overall technical DiSSCo Architecture
- Refine the DiSSCo technical architecture for digital specimen data and provide a plan for implementation
- Integrate DiSSCo in the technical landscape of national, European and global biodiversity data providers

Work focus M 1-12:
- **D5.4 A best practice guide for semantic enhancement and improvement of semantic interoperability**
- **Towards D5.1 DiSSCo Knowledgebase for technical development**
- **Compilation of relevant data standards**
- **Development of MIDS and OpenDS**

Overall progress WP 1, 3, 5, 6

- All WPs in the stream have made **excellent progress** already in the first year and are on track for deliverables.
- Minor deviations:
    - WP3 had a late start caused by Covid-19 measures which caused delays in MS3.1 and MS3.2, but MS3.1 has been delivered and the deliverable (D3.1) will probably be on time.
    - Work on the Knowledgebase in WP5 started earlier than originally planned and MS51 will likely be achieved earlier as planned (was:M18), although documentation of relevant building blocks may take time and is part of the milestone.
- **Risks:**
    - WP6 has no early milestones and therefore difficult to assess whether it is on track. T6.2 makes good progress, progress in T6.1 is less clear.
    - New Covid-19 measures may affect the progress

Results from M1-12

700+ user stories!

- MS 1.1: Corpus of life science user stories and use cases compiled
- MS 1.2: Corpus of Earth science user stories and use cases compiled
- MS 3.1: Improving Digital Capability – Case Studies & Analysis
- (MS 5.1: Functional technical implementation of DiSSCo Knowledgebase) – MVP available
- MS 5.5: Compilation of relevant data standards

- D 5.4: best practice guide for semantic enhancement and improvement of semantic interoperability

DiSSCo Knowledgebase MVP

AHM 1 S&T WP Highlights

WP1

- T1.1 and T1.2 **Use cases and User stories**
  - Initial progress made with grouping the stories, but careful thought is needed to make the grouping useful for prioritization and implementation
- T1.3 and T1.4 **Criteria for prioritization of digitization and indicators of socioeconomic impact**
  - Excellent presentations by Rui Figueira and Elsa Fontainha
  - Quantification of indicators is the major challenge
  - Will lead on compiling performance indicators throughout the project.
  - Sub-milestones defined

WP3

- T 3.2 **Collate and implement best practices of institutional data mobilization Kick-off**
  - Relations with T1.3, 6.1 and Synthesys 2.1
  - Needs discussion with WP5 on how Knowledgebase can be used consistently in this task
  - Build on ICEDIG outputs and liaison with iDigBio
- T3.1 and T3.3 **Competencies and capabilities – emerging thoughts**
  - T3.1: Digital maturity in organisations is separate to individual competencies and requires more work
  - T3.3: Need to understand and recognise incentivisation for secondment.

AHM 1 S&T WP Highlights (2)

WP5

- T 5.1 **DiSSCo Knowledgebase**
    - Presentation of current version, feedback and requirements by project partners discussed
    - DiSSCo Github to use for testing and further feedback

WP6

- T6.2 **OpenDS** (2 sessions)
    - Presentations and discussion on recent progress of the development of the Open Digital Specimen specification providing the fundamental data model for the DiSSCO Digital Architecture
    - A DO infrastructure brings opportunities for linking services (e.g. from GBIF) to be incorporated in different workflows but guidelines are needed.
    - Need discussion how CMS should synchronize with digital object stores

AHM 1 S&T Stream session

how to improve our internal processes in our organizations
to optimize our contribution to DiSSCo and the project

Possible success factors:
- In-house ICT/informatics as a means of achieving daily interactions between ICT and science
- Need for shared/common language, both scientific and technical
- Mixing up teams - putting them in different buildings/floors perpetuates silos
- Integrated view of digital across the institutions - articulate the business case
- Driving the agenda forward versus providing a service support function to researchers.

Need to think about the wide scope of informatics teams:
- Innovate/research (why should such teams have their own scientific/research agenda?)
- Continuously improve existing systems
- Making the case for change
- Communication/training
- Thinking through implications of organisational changes

Involvement of smaller institutions:
- Make it super-easy to adopt at the entry level
- Develop in a country-oriented way that allows all institutions to benefit

Reflections on AHM1

- Well received keynote talk from Prof. Isabel Sousa Pinto
- The stream session and WP sessions were well attended
- The sessions were very effective in making progress and having active discussion
- While dependencies and potential overlaps seem no big issue within the stream, AHM1 reveals a few areas in WPs between the stream that need further attention:
  - Performance and Impact Indicators developed in T1.4, T3.2.4, T4.1.1 and T6.4.4
  - Similar information needs for WP 2, WP3, WP5, WP7, WP8 which may required coordinated action for e.g. questionnaires and require appropriate contacts identified in the institutions.

Session Task 1.1 and 1.2 "Use cases and user stories"

Session outcomes:

Goal: Preparation of condensed list of functional demand

- Working session: almost 50 participants
  - breakout groups for certain (groups of) use categories
  - dedicated work on prepared tables listing requirements from use cases and user studies (field "For this I need")
- preliminary categories for grouping of listed requirements
- collections of questions and issues, discussion
- progress towards condensed list of functional demands as recommendation for WP5 & WP6

Session Task 1.1 and 1.2 "Use cases and user stories"

Next steps:

- add IDs to table to allow linking functional demands to specific use cases
- prepare condensed list of functional demands (after de-duplication, specification)
- agree on common vocabulary for categories of functional demands
- grouping of functional demands and analysis
- consultation with WP5 and WP6 (to specify which kind of recommendations are most useful)
- another joint meeting of Tasks 1.1 and 1.2 (probably end of February)
- Deliverables for Task 1.1 (Life sciences) and 1.2 (Earth sciences) due by April 2021

1. Session: WP 1, Task 1.1 and 1.2 "Use cases and user stories"

January 19, 2021 | Convener: Aino Juslén & Mareike Petersen
                      Co-Convener: Heli Fitzgerald & Sabine von Mering

Almost 50 participants attended the joint working session of Task 1.1 and 1.2 of WP1. Participants included members of all institutions actively involved in both tasks but also members from other work packages and the Coordination and Support Office (CSO).

List of participating institutions:
Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin, Germany
Cardiff University, UK
Finnish Museum of Natural History, Luomus, University of Helsinki, Finland
Meise Botanic Garden, Belgium
Museum für Naturkunde Berlin, Germany
Musée national d'histoire naturelle, Luxembourg
Muséum national d'histoire naturelle, Paris, France
Natural History Museum London, UK
Natural History Museum, University of Florence, Italy
Naturalis Biodiversity Center Leiden, Netherlands
Naturhistorisches Museum Wien, Austria
Royal Belgian Institute of Natural Sciences (RBINS), Brussels, Belgium
Staatliches Museum für Naturkunde Stuttgart (SMNS), Germany
University of Copenhagen, Denmark
University of Lisbon (ULisboa), Portugal
University of Tartu, Estonia

2. Main discussion highlights and decisions

The aim of the session was to produce a condensed list of functional demands, i.e. to add to, specify and condense the prepared collections of use cases and user stories. Also, the list should be prepared for further analysis as a basis for the upcoming deliverables D1.1 and D1.2, the reports on use cases and user stories for Life sciences and Earth Sciences with recommendations to WP5 and WP6. Both Deliverables are due in April 2021.

During the working session, participants were split up into four breakout groups focussing on different groups of use categories (breakout group I: Research, II: Collection Management, III: Technical support (IT & IM) and Policy, III: Education (academic & non-academic), Industry and External (Media & empowerment initiatives). This allowed dedicated work on prepared tables listing requirements from use cases and user studies from the Life and Earth Sciences as compiled for the Milestone reports MS1.1 and MS1.2 (https://dissco.teamwork.com/#/files/9103022). The use cases and user stories were captured in the "epic story" format (As a… I want to… so that I can… and for this I need…). The field "For this I need" contains requirements and (functional) demands that need to be sorted, deduplicated, grouped and specified to define precise functional demands.

Session outcomes:
- Some user needs may have two or more functional demands.
- Many user needs are duplicated taking into account the different use categories/user groups and need to be merged.
- There are some user needs which may not be the direct responsibility of DiSSCo but other infrastructure and which still may need to be considered - eg species' information.
- Subsequent checks are needed whether some functional demands are already covered by e.g. GBIF.
- Industrial use cases are sometimes difficult, e.g. consumer vs. digitizing company.
- School curricula would be an important use case (evidence from PISA that scientific collections are important), but difficult to deal with (different curricula even in one country, more throughout Europe)
- Some user needs are too vague to define concrete requirements (e.g. from category education) and need to be specified.
- Based on feedback from the group it was decided that IDs should be added to link the functional demands to the original use cases and user stories. An internal GitHub repository might be used to store the table(s) with all use cases/user stories (incl. IDs to link with functional demands), to work on the grouping and further analysis and to collect issues.
- Several groups suggested that it would be useful to characterize and group the functional demands in categories before working on the list of functional demands.
- Preliminary categories for grouping the listed requirements and demands are useful but need further refinement in a dedicated meeting of the task groups.
- A meeting with members of the WP5 and WP6 (also representing the Technical team) should evaluate if the grouping is useful for the further design of the DiSSCo RI.
  - Before this we need to gain a better understanding of what we have, e.g. are there "unexpected" use cases or functional demands that have not been considered so far?
- Additional interviews with representatives of different categories might help to fill gaps in the current list of user stories/use cases for further analysis and to gain more detailed information on functional requirements/needs of users.

Collected **functional demands** mentioned several times:
- Advanced search, harvest and import functionality
- Annotation system or Unified Curation and annotation services = UCAS
- Collection's level descriptions (in contrast to specimen level)
- Comprehensive collection catalogue (for digitized and non-digitized collections)
- Comprehensive metadata (incl. institutional)
- Dashboard (e.g. collection descriptions and digitisation level)
- Data standards, standards for linking
- Dynamic maps (incl. geographic query functions)
- Institutional metadata
- PID system (e.g. for digital specimens)
- Record system following metrics for the use of collections (publications, outreach, training)
- User-friendly and comprehensive portal
- Functionality for institutions to maintain data regularly updated or to automate data updates where possible
- Clear guidance on specimen citation in literature - tracking PIDs in literature

- Language issues need to be addressed (e.g. for institutions names), add language tag on certain fields (i.e. free text fields)

A more complete list of requirements and functional demands resulting from the use cases and user stories was compiled in a shared table where further work can be jointly done by the task partners.

Preliminary vocabulary for grouping the functional demands:
(starting point for further work)
- Tools
  o Tools for (data) analysis
  o Data discovery
  o Annotation
  o Download
  o Reporting & statistics
  o Documentation
- External services (outside collections community)
  o Operating with internal data
  o Operating with external data
- Procedures
- Data
- Metadata
- Physical access
- Images
  o Digital representation of specimen
  o 3D images
  o Label
  o etc.
- Standards (incl. taxonomic backbone, geological thesaurus)

3. Next steps

- Add IDs to table to allow linking functional demands to specific use cases
- Prepare and update condensed list of functional demands (after deduplication, specification)
- Agree on common vocabulary for categories of functional demands and define rules for categorising (how many categories per functional demand? Is a supercategory or subcategory structure required?)
- Grouping of functional demands and analysis
- Consultation with WP5 and WP6 to specify necessary feedback and recommendations, i.e. to discuss the level of detail needed
- Additional interviews with representatives of different use categories to integrate their stories in the analysis
- Another joint meeting of Tasks 1.1 and 1.2 (probably end of February)
- Deliverables for Task 1.1 (Life sciences) and 1.2 (Earth sciences) due by 1 April 2021

4. Conclusions

During the joint work session considerable progress was made towards a condensed list of functional demands. These will be further developed into recommendations for WP5 & WP6 and thus will help to build the DiSSCo Research Infrastructure.

5. Reference documentation

Link to Milestone report MS1.1 & MS1.2: https://dissco.teamwork.com/#/files/9103022

Task 1.3. Establish relevant criteria to identify a prioritisation model for digitisation

- ICEDIG's D2.1 is an excellent starting point
- Institutional digitisation plans should be analyzed.
- Also look at SYNTHESYS+'s Virtual Access and non-European projects.
- Multiple drivers of digitisation must be considered, incl.
  - National vs. institutional priorities, funding opportunities, cost, technical constraints …
- The T 1.3 deliverable may be developed into an online wiki and/or be added to DiSSCo's Knowledgebase
- **NEXT STEPS: Obtain institutional digitisation plans**
- **Explore the "decision tree" developed by ICEDIG**
- **Bimonthly task meetings; first meeting primo March**
- **Hire part-time staff for T 1.3 in Copenhagen.**

Task 1.4: Develop indicators of socioeconomic impact
Purpose of the session:
- Kick-off the task
- Review the main existing frameworks for indicators of socio-economic impact of Ris
- Identify challenges and particularities of socio-economic impact indicators for RIs in the Environment
- Define next steps

Review of the existing frameworks and previous work

Report by OECD:
- reference model
- challenges of Ris
- 25 core + 33 standard indicators (name., description, data)

OECD. Reference framework for assessing the scientific and socio-economic impact of research infrastructures. (2019)
doi:https://doi.org/10.1787/3ffee43b-en

Report by ESFRI Working Group (KPIs):
- 21 quantitative indicators
- qualitative indicators
- sheet will full description of each indicator (good template)
Report of the ESFRI Working Group on monitoring RIs performance | www.esfri.eu.
https://www.esfri.eu/latest-esfri-news/report-esfri-working-group-monitoring-ris-performance (2019)

Requirements for defining socio-economic impact indicators
Be innovative in integrating environmental and social factors
Review other RIs + DISSCO proposal + scientific reports

Next steps:
- review and compile existing indicators, assessing its adjustment to DISSCO (3 months)
- Interact with WPs leaders and Streams leaders to evaluate adequacy/needs for implementation/other (6 month)
- first meeting - 2nd half of February

Task 3.2 AHM Meeting Agenda & Notes

Background

The formal title for Task 3.2 is "Collate, refine and implement best practices for data mobilisation at the institutional level to develop the DiSSCo plan for data mobilisation and curation pipelines".

The full task description is provided in the linked planning document along with the subtask descriptions.

Participants

Laurence Livermore	Alex Hardisty	Frederik Berger
Helen Hardy	Elspeth Haston (had to leave	Sofie De Smedt
Mil de Reus	at approx 12.50 GMT)	Tina Loo
Anton Güntsch	Josh Humphries	Pedro Arsenio
Esko Piirainen	Maarten Trekels	Roger Hyam
Mathias Dillen	Maria Joao Santos	Tania Walisch
Patricia Mergen	Quentin Groom	Elsa Fontainha
Patrick Semal	Vince Smith	Kari Lahti
Heli Fitzgerald	Serge Scory	RosaRosa (Unifi)
Ann Bogaerts	Sarah Rossi de Gasperis	Rob Cubey
Anne Koivunen	Wesley Tack	Carole Paleco (RBINS)
Wouter Addink	Judite Alves	Tania Walisch (MNHNL)
Ana Casino	Ville-Matti Riihikoski	

Draft Agenda

- Task Overview
- Discussion
    - Current state of institutional digitisation
    - General state of natural science collections digitisation
    - Task logistics

Minutes

Questions and discussion

Questions shared in chat:
What are you currently digitising (or planning to do when you return on-site)?
Has COVID affected your institution's plan or outlook on digitisation?
How is your institution monitoring digitisation at the moment?
What is everyone digitising (or planning to digitise) beyond the well-tested and fairly well-published workflows (pinned insects, microscope slides, herbarium sheets)?
Do you have pre-digitisation protocols? How do you assess and prioritise collections?

Do you have well-documented internal protocols for your workflows? How much of them can be generalised? Do you have documented hardware and software? How much post-processing and temporary (logistics/operational) metadata capture do you use?

What are your pain points/constraints in digitising data? What do you feel could most easily be improved (considering costs and benefits)

How do you publish and share digitised data?

What is the current state of institutional digitisation?

LL asked for a quick summary of key projects/programmes now (or on return to Museums depending on Covid restrictions).

Luomus is not too affected by Covid restrictions.

A lot of work to digitise insect specimens.

Use of conveyor belts for imaging.

Also transcribing.

Plan to acquire a CT scanner this year, and two more imaging stations for lichens and mosses - photographs plus OCR to read labels (new for Luomus - this collection has more typed labels so should be more straightforward).

Meise in 2nd mass digitisation project, to digitise 1.2m specimens which are the general herbarium using Picturae. Had to stop last March but then restarted after a month with Covid precautions such as barriers - brought rate down from 4,000 per day to 3,000. Imaging now expected to continue until May. Transcription keeps proceeding using the team in Surinam.

In-house digitisation of micro-algae also stopped but now restarting - only allowed to access the collection one day a week at present. Transcription did continue at home, using citizen science platform from first digitisation project on Belgian collections. As a citizen Science project, this saw high uptake while people were at home during Coronavirus.

Berlin currently is not able to digitise on site so focusing on transcription. Running a Zooniverse project on bee labels (a large one). When back on site will start up in entomology - tendering for a system with throughput of 5,000 specimens per day similar to a herbarium approach. Also work on dry invertebrates (mollusca and fossil inverts) - have a new camera system, not yet measured but expected to speed up in house digitisation of this material.

Also developing 3D and High res imaging.

CT scanner installed and expected to be operational next month.

RBG Edinburgh most digitisation in house (core funded). Recently worked with a local company to replace light boxes. Also purchased new cameras to go with these.

A major migration to Specify from old CMS - working to develop data entry tool for this for digitisation.

Crowdsourcing projects with DigiVol and also working with Zooniverse

CMS migration slowed work as much or more than lockdown.

Work on MIDS(minimal information about digital specimens standard) with CETAF et al - keeping it simple on imaging (is there any image or not?).

This task will need to work with the CETAF Digitisation Working Group. A key connection.

Lisbon - research infrastructure in Portugal for biodiversity data. Some extra funding in last 3 years for digitisation in collections, including herbarium, zoology and a little in geological collections. A lot still to be done in standardising workflows - different collections do own programmes at present. Funds running down and not sure how will fund digitisation in future (affects different departments/institutions in Lisbon) - makes it hard to keep any kind of 'mass digitisation' rhythm.

NHM mostly remote working last spring-summer, now focusing on transcription. When back on site will be focusing on Synthesys+ Virtual Access projects - freshwater flies and bats. First zoological project with the bats. Also looking at workflows for carpological herbarium collections; and maybe fish depending on staff numbers.

RBINS: Micro CT is running normally because it is a separated building with possible social distancing. The digitization on site is allowed 1 day a week with homework for the treatment of the data using remote applications like "teamviewer". The manual data and metadata encoding continue normally but as homework with quality control of the encoded data. Adaptation of the Collection Management system to CETAF Stable identifier and IIIF compliance for the image server (common work with Africa Museum Tervuren)

Is anyone mass digitising anything other than slides, herbarium sheets and insects who hasn't already spoken?
NB there isn't a convenient definition of mass digitisation (although there is one from ICEDIG - shared in meeting chat by Alex Hardisty). Just looking at cost or throughput can be misleading but is a broad guide.
None raised anything they felt would be considered mass digitisation not already mentioned.

NHM mentioned constraints on getting data into CMS at mass scale. Question to the group for other blockers/pain points?
Berlin - hard to find funding to scale up or to hire people on short term contracts with relatively low sector pay. Work with industry is also challenging.
Luomus - In a relatively fortunate position, but roles can be challenging e.g. where staff have other responsibilities as well as digitisation. Can take time to resolve any hardware or software issues.
RBGE - Issues with importing data into CMS, particularly in the context of migration. Relates to the nomenclatural backbone (would like to import International Plant Names Index). Have electronic data on many historical collections - couldn't previously import this but that may become easier with the new system. Similarly for crowdsourced data.

Pat Mergen - **when DiSSCo is a legal entity would it be helpful to issue tenders above institutional level? Could support outsourcing. A question for a later date.**
Still requires funding of course. Examples do exist of individual and small groups of organisations tendering. (https://ec.europa.eu/digital-single-market/en/pre-commercial-procurement) linked to Task 4.4 of DiSSCo Prepare

Has anyone used or adapted a documented workflow (e.g. from a paper)?
How widely are barcodes used and should that be assumed best practice for DiSSCo?
WA - yes for newly generated digital material, would advocate adding barcodes as part of digitisation process.

Meise - found papers useful, particularly on imaging quality and data standards etc. Helped in preparing tender documentation. Also looking at NHM microscope slide workflow. Standards includes broader ones than our community e.g. TIF standard or colour quality control etc. Sector-specific standards and points tend to be more known or accessible already? Could use this task to complete write-up and publication of some workflows from ICEDIG that were not completed.

There was an ICEDIG deliverable on quality control for imaging.

LInks to some material from prior projects are in the planning document.

This task will need to work with the knowledge base. Need to ensure a consolidated collection within that e.g. for SoPs. Publication is a bit separate.

Knowledgebase will have some metadata and API.

IDigBio also has resources - how could we help to bring resources like this together and make them more searchable? And have data e.g. on when procedures were last updated or similar.

ACTION - review key IDigBio resources as part of this task and provide a way to access the most useful?

HH - liaise with IDigBio on this? They have a very wide base of useful material e.g. they have some relevant to WP3 as well]

ACTION - Get in touch with WP5 on how Knowledge Base can be used consistently by this WP

WA - where is the crossover from having best practices to more action to encourage their use?

Probably not in scope for this task - a recommendation - but could be part of work e.g. in MIDS? Could look to set more requirements? Perhaps this goes beyond the Prepare phase of the project? A transition towards the end?

What about recommendations on the process for conversion of best practices etc to use? (even if the actual conversion is premature)

A lot depends on the granularity of SoPs - often reasons to deviate from a detailed process. Also a constant change - need to reflect the need to accommodate innovation in a sustainable way. [HH note - does the IDigBio wiki do this at all?]

Task logistics

How frequently should the group meet? Fortnightly agreed. **ACTION LL to schedule**

Cetaf Digitisation Working Group and ISTC will be very important.

How open is the MIDS discussion? Takes place in TDWG task group chaired by Alex Hardisty and Elspeth Haston - can join the email group on request. Also the CETAF DWG who feed in best practice suggestions etc. Open to CETAF members.

There is a Github repository here: https://github.com/tdwg/mids

Alliance for Biodiversity Knowledge and others consulting from Feb about digital specimens and extended specimens - one of 5 strands of this will be about mobilising FAIR specimen data. Alex H co-moderating that as is Wouter Addink and Barb Tiers (New York).

Alex Hardisty chat comments:

Alliance For Bio blog post on forthcoming global consultation: https://www.allianceforbio.org/post/converging-digital-and-extended-specimens-towards-a-global-specification-for-data-integration .

Background document on the consultation: https://docs.google.com/document/d/1mtjLD7Zpf73apLajW8qStA3U0Gpz2ZGazZkjN2Vp3Ew/edit

Strand 1 in the consultation, on "digitizing/mobilizing FAIR data for specimens" is relevant for task 3.2 participants. Contributions would be welcome. Consultation is likely to be open from 16th February to 5th March.

TDWG Task Group on MIDS: https://github.com/tdwg/mids and the current draft of the specification: https://github.com/tdwg/mids/blob/working-draft/current-draft/MIDS-definition-v0.12-03Nov2020.md . TG MIDS aims to meet monthly a few days after the CETAF Digitisation WG meeting, which also aims to meet monthly. https://cetafdigitization.biowikifarm.net/cdig/

Luomus have volunteered to work on standardised ETL, Meise on pre-digitisation curation, and NHM on SoPs.

Who could lead subtask on digitisation monitoring? Edinburgh willing.

NHM will review existing material from IDigBio and consider gap analysis across this, ICEDIG etc on digitisation workflows/SoPs. Areas where workflows don't yet exist or have a lot of variance. Probably also extends to quality assurance.

Luomus - would like to understand the subtask goals more clearly around ETL and start with a discussion about this.

CMS import as discussed above is of interest but also other aspects.

Esko - The overall architecture is relevant e.g. links to other systems/platforms etc

Some links to other work packages. WA - some work on CMS system requirements Task 6.1 led by MfN and WP8 task on OCR.

ACTION to coordinate this sub task with these or other relevant tasks.

Need a clear separation in steps of digitisation process - e.g. so that OCR could take place in a distinct tech 'pipeline'?

Pre-digitisation and monitoring similarly will need a gap analysis / analysis of previous work, and discussion with institutions.

Check SYNTHESYS Plus task 2.1 on policies, which addresses also needs in terms of Pre and post digitization notably in the framework of Virtual Access and future Digitization on Demand within DiSSCo.

Anyone who is closely involved in digitisation or in technical aspects of data mobilisation please do make themselves known.

Work to create a list of principle contacts for key topics is underway (by CSO? WA aware of this)

Jose Alonso - There is a helpdesk aspect around what should be provided by the helpdesk - e.g. could include documentation?

Would this include the handbook? Not started yet but likely.

ALL HAND MEETING - AHM1
T2.1 Wrap-up

18 January 2021

Ana Casino - CETAF

DiSSCo Prepare – Training Strategy

Task T2.1 Leader

Funded by the Horizon 2020
Framework Programme
of the European Union

T2.1 Training Strategy – KoM

Friday, 22 January (10h45-12h15) – Ana Casino (CETAF)

D2.1 Training Strategy recommendations (Dec'22)

- Agreement on shared understanding of the task

- Distribution of work in sub-tasks and along a propsoed workplan (18PMs)
- identification of contributors and expertise brought into the task (6 partners)
- Alignment with other projects' outcomes and sister initiatives:
    * DEST, as a platform for allocating
    * SYNTHESYS+ D2.3 – Training catalogue / roadmap (April'21)
    * T3.1– Competences and skills

Upcoming deadline: MS2.1 "Recommendations on suitable training mechanisms" (Jan'22)

Meetings: monthly meetings, and a very close interconnection with WP3 and SNTHESYS+ T2.3

Challenges to address:

- Realm to cover and related topics (platforms, service business model, multilingual…)

DPP AHM1 T2.1 Training Strategy session

Date: Friday January 22nd, 2021

Time: from 10.45 to 12.15 CET

Organizers: Ana Casino, Judite Alves and Marie-Laure Kamatali

Presenters:

- Hugo de Boer (NHM-UIO)
- Carole Paleco (RBINS) and Magalie Castelin (MNHN)
- Helen Hardy (NHM)

Reference Documents:
- T2.1 Work plan suggestion (word doc format):
  https://docs.google.com/document/d/1X3QDoNSCg1JsHmPlCi1kB6LKmZ-1cjU01k5famO5JZI/edit#
- T2.1 Work plan suggestion (excel sheet format):
  https://docs.google.com/spreadsheets/d/16puw4_EtYuynxiyCrkdxWZnuOC5LtTXmMg3HiGK_BGQ/edit#gid=213675108

Resources:
- ICEDIG deliverable D8.1 Conceptual design blueprint for the DiSSCo digitization infrastructure
  :https://riojournal.com/article/54280/
    - To do: read recommendations n74 to n100 which cover training needs.

Agenda

Time	Topic	Lead
10.45-10.50	Welcome	AC
10.50-10.55	Aims ● Common understanding of the task ● Distribution of work and responsibilities in accordance to partners' expertise ● Establishment of a regular working channel	AC
10.55-11.20	Tour de table of representatives	All
11.20-11.30	Presentation of the T2.1 "Training Strategy" ● Context within WP2 ● Objectives ● Interactions with other DiSSCo linked/associated/related actions ● Expected Outcomes	AC/ JA /MJF

	● Timeline	
11.30-11.45	Alignment with sister initiatives and previous developments ● DiSSCo PPP WP3 - T3.1 "Improve digital skills and competencies across DiSSCo facilities"- HH (NHM London) ● Contributions to/from DEST- Distributes European School of Taxonomy - HdB (NHM-UIO) ● Inputs from SYNTHESYS+ T2.3 - CP (RBINS Brussels)+MC (MNHN Paris)	AC
11.45-12.05	Work plan draft ● Task/Subtasks ● Content ● Contributions / Responsible person ● Timeline ● Open discussion	MLK
12.05-12.10	Identify next steps ● Regular meetings ● Online work Mechanisms to collect information from DiSSCo partners	AC
12.10-12.15	AOB	AC

Session Type: working meeting to kick off work in T2.1.

Notes Taker: Céline Cassarino and Laura Tilley (CETAF)

Participants:

Rosarosa Manca (NH Museum, University of Florence)

Judite Alves (Museu Nacional de História Natural e da Ciência, ULisboa)

Sarah Rossi de Gasperis (NHM, University of Florence)

Helen Hardy (NHM)

Sharif Islam (Naturalis/DiSSCo CSO)

Anne-Sophie Archambeau (MNHN/IRD, Paris)

Magalie Castelin (MNHN, Paris)

Anne Koivunen (Luomus)

Quentin Groom (Meise)

Lorenzo Cecchi (NHM, UniFI)

Patricia Mergen (MeiseBG)

Niels Raes (Naturalis)

Piotr Tykarski (UW, as an observer, not involved in the WP2)

Luca Bartolozzi (NHM, UniFI)

Pedro Arsénio (ULisboa)

Bruno Ribeiro (ULISBOA)

Carole Paleco (RBINS)

Elsa Fontainha (ULISBOA)

Frederick Berger (MfN)

Hugo de Boer (NHM-UIO)

Jose Alonso (Naturalis)

Luca Belluci (NHM,UniFi)

Maarten Trekels (Meise)

Maria Joao Santos (ULISBOA)

Nele Van der Schueren

Pedro Arsenio (ULisboa)

Peter Giere (MfN)

NOTES:

Aims of meeting
- 3 major objectives
- Common understanding of what this task
- Regular working channel

Common understanding:
- Training strategy will compile the key components and recommendations and how to the training actions as a service to the science community:
- Key outcomes: two milestones and one deliverable: - recommendation of suitable (December 2021), - landscape analysis of best practices (defined and delivered in the training strategy) – due July 2022
- Timeline starts today and up to the 2 years
- It shall include the community capacity building
- We need to identify the portals and channels in which the training will run: what is the business model behind the training programme who is going to give the training.
  - We need to first identify the needs and how we ensure that they are well covered, which are the gaps how do we embed need to take into account the diversity of institutions
  - Also to consider the profile of the training staff, standards protocols etc
  - How can we enhance performance as users of the DiSSCo platform. So we can take the benefit of the structure.

Linked outcomes from other DiSSCo linked projects:

- ICEDIG did provide a design blueprint that had recommendations please go through this 74 - 100 - for seeing the needs etc.
- SYNTHESYS+ will provide the road map and schema. Led by MNHM and RBINS
- DPP T3.1. Consultation of skills and competencies we should all have a look at dashboard under WP based on the above analysis will be a tool that we can embed results.

Partners and involved actors from each institution: new names need to be added as contact points and people that will be hands-on during the duration of the task.

Questions
Patricia Mergen_Meise: is it possible to invite colleagues from T2.3 and Gbif but not in institution can participate - Ana says there more resources the better they more than welcome. The only thing that has to be on volunteer basis know reimbursement. Please let us know so we can invite them.

WP2 Context by Judite Alves (ULIBOA)

WP2 is about the dissco community human resources and policies behind , training and to provide support. The mission is to make sure the users and data providers have the capabilities and skills

Key objectives addressed in each task
- Training strategy (T2.1)
- To design user support trainer support services - for dissco services (T2.2)
- Human resources for DiSSCo so they are used efficiently (T2.3)

WP2 objectives are interrelated, they have the dissco community at the center, the human resources will frame the training strategy on the other hand the the help desk will bring the training strategy to the community we will need further training for the help desk. We will enhance the community towards the use of the infrastructure. We will build on linked projects including MOBILISE, and platforms DEST and biotalent.

These are related also with the helpdesk also we are dependant on other work packages lik 3.3. 3.1 wp7 and wp 6 (provide info on services) and the documents that have to be provided in the wp. We are not the only one dealing with these challenges.

Comments:
AC: two folded profiles think that it is essential without the competencies we may not be able to fully use DiSSCo thus we need to upgrade the competencies of our tasks. Agrees that it is fundamental and

to develop the right catalogue for the services this is a challenge. We need to combine the landscape complexity and the training needs of the community

JA - the landscape is very difficult to address and is better complexified by the heterogeneity (different levels of skill, languages) thus the training task is complex it is good we have different partners and bring different country views.

AC this will be quite related to WP 3 but also their linkage to WP7 with the policies and the variety of our institutions to access DiSSCo. Thus the needs of training may defer a lot also the domains we need to tackle the linkage between science and finance (WP4) we need to understand the process of how dissco will run there are a lot of linkages.

ALIGNMENT with SISTER INITIATIVES

   a. **DiSSCo PPP WP3 - T3.1 "Improve digital skills and competencies across DiSSCo facilities" by Helen Hardy (NHM London).**

The task T3.1 is aiming to improve digital skills and competencies. Capacity building is a key element to support DiSSCo related services.

Ongoing outcomes:
   - **Report delivered** that cover case studies (MS 3.1 -Dec 2020) 'consultation and analysis')
   - **Ongoing discussions around the implementation of a tool (dashboard)** that could be useful to connect and map capacity building skills and competencies around data (MS 3.2 - April 2021)
   - Update & identification of gaps + reorientation for actions thanks to the first report on 'case studies' (MS 3.1) - MS 3.3 (May 2021)
   - 'Digital & data competencies framework - final tools and recommendations for the deliverable (MS 3.4 - July 2021)

Areas of exploration (from sub tasks):
   - **Competency frameworks for individuals**
Looking and mapping digital and data expertises and also to management and leadership of people and projects that are involving NSCs - hope to feed in by Summer 2021

Discussion: should we develop **a self-assessment tool for DiSSCo** so to understand who are the data-related experts, what training they are offering among the partner organisations?- they could provide that training as an internal consultancy. **HH:** Could DISSCO award partners to provide training? - this requires a certain incentive model.

- **Organisational change to support digital capacity/capability**

Including communication, policy, governance, outreach, and organisational structure, and developing towards customised recommendations

- **Digital capacity data**

E.g. surveys: how data collection could be automated and/or data be made more machine readable

AC: the dashboard could be the starting point to identify the needs. To be considered in the business model: who can be the trainers internally, ie? how to credit the work done by the trainers? part of the recommendations. It can't be a volunteer basis. There are few points to be discussed, we need to collaborate, having you as a guest (HH) at meetings would be useful and vice versa.

AC:

- to which extent the dashboard (3.1) could also be the starting point for us to feed and identify the needs to be covered.
- Business model to be implemented. Who can be the trainers internally?
- How to credit our own trainers? This is part of the recommendations: we need to study and analyze the possibility to provide external resources, it can not be on a voluntary basis if we envisage long term developments

b. **SYNTHESYS T2.3 by Carole Paleco (RBINS) and Magalie Castelin (MNHN)**

T2.3 dedicated to developing focus training activities (presented by CP)
Key question: what is it to develop training activities?
Task involves 13 partners, 11 plus GBIF, DEST TDWG.
One deliverable on catalogue of training and one milestone in Feb, workshop held in paris. To define the training program.

First phase
Objectives: to track transnation and VA and support the community to get digital competencies. Identification of training mechanisms, specific training modules, key training areas needed.
Organisation of train the trainer events.
Collaboration extended with DEST and planning of a workshop in Paris that sought to define prioritisation needs.

Outcomes of workshop (presented by MC):

- A list of training shared on a Google Drive Excel File
- Advertising campaign on CETAFF website
- A milestone uploaded on teamwork in SYNTHESYS+ project. (MS26 Report on Training Progr. Definition, 15pp)
- Data regarding location, length of training modules (among others) was collected, which was then categorized in key training areas. This allowed identifying some training gaps. We are still in the middle of this analysis (*Description of the training landscape & current training mechanism).*

Second phase (February 2020)

Objectives:
- how to expend the catalogue and refine the gaps identified
- give recommendation for the development of a proactive, efficient, and evolving DiSSCo Training Programme**.**
- How to obtain recommendation for a effective training plan.

Currently: analysis on IRL mapping and developing the deliverable D2.3.
Second analysis focused on each IRL area and studied all the goals with in each dimension to define sub categories, DPP WP4 cost book was used as it's a comprehensive list of skills and tasks needed for DiSSCO.

2 examples defined:

- (1) a subcategory of the IRL TECHNOLOGICAL: *IT data management*: gather activities around security of data (the manipulation, protection, storage, or safety of data.)
- (2) a subcategory of the IRL TECHNOLOGICAL:*Development* all the skills for designing software, database (e.g., data architecture, data carpentry – i.e., development of ELViS))

Other actions included:
- The idea of exercise to try and catch potential needs for each phase of dissco and that can evolve with the changing needs.
- Gap analysis: table of all training from institutes and tried to match every training to all sub categories.
  - the subcategory *E-monitoring* might be a gap (Digital management services, e.g., e-service, helpdesk, mediation services) for both administrators and users.

Current actions:

- Working on draft of the catalogue in april, need feedback from partners
- Situation analysis of existing trainings
- DiSSCO training needs assessment
- Recommendations for training programme.
- The training catalogue will be on DEST currently working to see how to upload the data.

Comments:

AC: importance of having the connection between training and the DiSSCo dimension. The training strategy (T2.1) needs to cover all the needs that our staff need to have, and how the work in SYNTHESYS+ will facilitate this.

HdB: questions about the catalogue wondering if analysis of the distribution and access (from language to availability) was made within the network we offer lot of courses but are not available due to language differences

- MC - yes we have tried to do this - including fees and certificates.
- Hugo- we see in dest there is a distinction between data providers and supplies virtual courses how can we make them accessible
- CP: data providers were asked to specify language of training.
- NR: users are not represented in the presentation; the providers and external users. We see training from internal users.

AC: important work by SYNTEHSYS+ in detecting the gaps of existing training provision. In DPP we shall reach the gap how to ensure what is need from one or other perspectives.

NR: GBIF 2004 species distribution modeling and this increased the users something similar for dissco should be made.

   c. **DEST by Hugo de Boer (NHM-UIO)**

Brief update - its a research training school offers training opportunity to students etc from any nationality - focused on data users
Set up new website for the DEST training portfolio - available now
There is a limited portfolio of courses hope to increase the number when the courses from SYNTH and CETAF members use the platform to advertise the courses
How can we increase awareness to provide good courses? We need virtual courses, DEST small team waiting to hear from you to help organise courses or create a website page that advertise your course.

Reflections on DEST - come forth from work with other people in WP5 training capacity enhancement reading ICEDIG recommendations that covers all the possible needs other activities split between data suppliers and users. The SYNT+ analysis a good way to analyses see where the gaps are

For dissco to achieve enhancement goals could use dest - dest can be expanded current scope is too narrow. Implement training and enhancement require more e.e. Coordination and organisation to ensure all gaps are filled and we provide high quality training

TOUR DE TABLE: each participant introduces themselves and shares in what ways they will be able to contribute to the task.

Niels Raes_Naturalis: dutch gb node mobilising data to gbif , biodiversity informatics how to use GNIF data there data should be provided

Hugo de Boer_NHM-UIO: Oslo represents DESt research school for bio Scandi - working a while data capacity in favor creating training school in taxonomy in DiSSCO objectives. Extensive experience and capacity building and training

Carole Paleco_RBINS: involved in DEST - dedicated to future infrastructure experience from SYNTHESYS+3.

Magalie Castelin- MNHM: paris molecular invertebrates database loans registration of specimen and trainer, as part of DEST, in the big-data era.

Sarah Rossi de Gaspris_UniFi: recently research grant DPP working in Florence about insect collection digitisation.

Lorenos Cecchi_UniFI: to the first subtask to complete the panorama regarding institutions-

Rosa Manca_ UniFi: conservation scientist - digital catalogue in art museum-

Peter Giere_ MfN - embryological collection provides and organises training involved in SYNth and DiSSCO.

Anne-Sophie Archambeau_MNHN: manager of gbif France active trainer data mobilise organising and doing training go fair programme IRD new training developments

Luca Bartolozzi_UniFi: Palaoentology Florence important interesting

Piotr Tykarski_UW: GBIF training training biodata projects - pure taxoniy interested in the topic.

Pedro Arsenio_ULISBOA: new observer not directly in task as complementary landscape architect involved in Data provide for GBIF

Luca Bellucci_UniFi: collection manager and data manager - rosa and sara main contributor.

T2.1 WORK PLAN and TIMELINE PROPOSAL
Three main outcomes connected to four sub tasks defined in the grant agreement:
2021 will focus on:
- **MS2.1 Recommendations on suitable training mechanisms**
Subtask 1 - Compilation of needs for skills/competences
Subtask 2 - Identification of training platforms and providers

2022 will be the year the outcomes below are submitted:
- **MS2.2 Landscape analysis of BPs for training delivery completed**
Subtask 3: landscape analysis of best practices for training delivery.
- **D2.1 Training Strategy**
Subtask 4 - Integration of all training strategy elements, to jointly provide a final report with a recommendation for setting up the DiSSCo training.

Timeline aspect: the task will be informed by SYNTHESYS+ catalogue presented earlier, which has been postponed to April 2021. Idea is to review it together in May 2021 during a working meeting.

Next steps will be:
- CETAF will share work plan to all task partners to collect feedback by Jan 29th.
- Organize our collaboration: send out doodle poll to schedule working meetings.

Frequency of meeting: monthly as starting point and then according to needs.

ZOOM CHAT

10:45:33 From Marie-Laure K. to Everyone : Hi everyone, could you please write down your names under the participants section in the following document:
https://docs.google.com/document/d/1GUK1qfIcCNkYYkRtUscVsO_ywwMMPomqqYiFpPZ70NM/edit#
10:45:35 From Marie-Laure K. to Everyone : Many thanks
10:46:28 From Marie-Laure K. to Everyone : re sharing it for those that just entered the meeting:
https://docs.google.com/document/d/1GUK1qfIcCNkYYkRtUscVsO_ywwMMPomqqYiFpPZ70NM/edit#
11:02:26 From Luca Bartolozzi UNIFI to Everyone : several people are to be added in UniFI
11:03:06 From Luca Bartolozzi UNIFI to Everyone : Sarah, Rosarosa, Lorenzo, Luca Bellucci
11:03:14 From Marie-Laure K. to Everyone : hi Luca, is it the same list as the one you provided for TW?
11:03:20 From Luca Bartolozzi UNIFI to Everyone : yes
11:03:37 From Marie-Laure K. to Everyone : we will amend the list of contact points
11:03:46 From Marie-Laure K. to Everyone : thanks for catching that
11:04:03 From Luca Bartolozzi UNIFI to Everyone : all the team is involved, even if most of the work wlil be done by Sarah and Rosarosa in 2021
11:04:27 From anne-sophie archambeau to Everyone : Sounds good to me
11:05:42 From anne-sophie archambeau to Everyone : I agree with Patricai
11:07:00 From Peter Giere to Everyone : I'll be involved for MfN
11:10:13 From Marie-Laure K. to Everyone : @Peter, well noted
11:41:53 From Patricia Mergen to Everyone : Ah ok.
11:42:27 From Patricia Mergen to Everyone : Sorry wrong window ...
11:51:32 From Marie-Laure K. to Ana Casino(Direct Message) : Hello Ana, the time is already at 11.50.
11:54:19 From Patricia Mergen to Everyone : COST Mobilise trainings also consulted
12:15:31 From l.cecchi@unifi.it to Everyone : Sorry, I just wish to add that I'm currently in touch with the large and very heterogeneous community of Italian Herbaria, where the need for training is quite significant...
12:15:50 From anne-sophie archambeau to Everyone : Sorry but I think it was all Ird will work e-trainings
12:18:01 From mariajuditealves to Everyone : Thank you Pedro

12:18:58 From mariajuditealves to Everyone : Thank you Bruno — I am counting on your close collaboration

12:26:21 From Helen Hardy to Everyone : I have to go now sorry - thanks for including me and I look forward to further discussions

12:26:32 From Pedro Arsénio to Everyone : Thank you all for a very interesting session

12:26:44 From Luca Bartolozzi UNIFI to Everyone : ciao to all!

12:27:09 From Luca Bartolozzi UNIFI to Everyone : non too often please...

WP 3 Task 3.1 and 3.3 – skills, competencies and distributed working

Task 3.1 & 3.3 meeting held Tues 19th – c.43 attendees

Agreements / outcomes:

- Overall, there is more support for a digital maturity/capability self-assessment tool for organisations and/or teams than for a DiSSCo competency framework (though tool might reflect key competencies)

- A good early kick-off discussion for 3.3 – physical secondment agreed less relevant now than piloting distributed development team.

Actions / next steps:

- Review additional materials / examples shared

- Further discussion with related WPs and tasks

- Consider CSAT example for self-assessment tools

- Further work on data sources

Open Issues:

- For 3.3 – how can we incentivise institutes to release time of specialist staff for greater good?

DiSSCo All Hands - Task 3.1 / 3.3

Competency and Capability

Monday 18th January 2021, 13.00-14.30 CET / 12.00 - 13.30 GMT

Notetaker(s):

Laurence Livermore (supported by Vince Smith and Patricia Mergen)

Purpose of session

This session aims to share learning from the Milestone 3.1 report *Improving Digital Capability - Case Studies and Analysis,* across both individual and organisational competencies and capability. Reflection on this will be an essential step to the next stages of work on this Task, including discussion of the most useful elements to progress for a 'dashboard' or set of tools. There will also be a short kick-off discussion for the related Task 3.3 about secondment and distributed working practices.

Attendees will be expected to contribute views about approaches used in their own institution (or other organisations they are aware of); thoughts on the examples in the report and the direction of travel these may indicate; and their feelings about what would be most useful and sustainable going forward. It would be helpful to at least skim the report in advance of the session - this is the key input for the session and is available on Teamwork under Files / Outcomes & Guidelines / Milestones.

13:00 - 13.10   Welcome and introductions

13:10 - 13:30   Presentation of key elements from the Milestone report and early thinking about these
Slides:
https://docs.google.com/presentation/d/1ZE-ZZFcI1sue1auhlgldNF8PUTe78r6omnaEFlxT4F8/edit?usp=sharing

13:30 - 14:00   Discussion of the learnings from the milestone and the next steps, including but not limited to:
- What kind of dashboard(s) or tools might be most useful in practice and why?
- Are there sources of data that the group have missed or should explore at the next stage? What is the best way to get further information about practices in DiSSCo institutions?
- What is most important to focus on among a competency framework; generic role profiles; capability self assessment tools for organisations; and/or ways to find organisations and individuals with relevant capabilities?

14:00 - 14:20 Task 3.3 - introduction from the task leader (Helen Hardy) and short discussion of early steps including possible survey or structured interviews about distributed and distanced working during Covid

14:20 - 14:30 Summary, next steps and close.

End-of-session outcomes will include session accomplishments / agreements; open issues; and actions / next steps. These will be shared at the Wrap-up session.

Attendees (A to Z by first name please!)

Peak 43

Helen Hardy (NHM) - Task leader Laurence Livermore (NHM Mil de Reus Dimitris Koureas Tina Loo Wouter Addink Alex Hardisty Alexandra Cartaxana Ana Casino (CETAF) - yes Ann Bogaerts Anne Koivunen Carole Paleco (RBINS) Claus Weiland Elsa Fontainha (ULISBOA) Ekso	Frederik Berger Gergely Babocsay Heli Fitzgerald Kari Lahti Lorenzo Cecchi (NH Museum, University of Florence) Maarten Trekels Mario Joao Fonseca Maria Judite Alves Marie-Laure Kamatali (CETAF) Matt Woodburn (NHM London) Mikko Heikkinen Patricia Mergen (Meise Botanic Garden) Patrick Semal (RBINS) Pedro Arsénio (ULisboa)	Peter Giere Peter Warth Pierre-Yves Gagnier Pieter Huybrechts (Meise Botanic Garden) Patrick Semal (RBINS) Quentin Groom Rosarosa Manca (NH Museum, University of Florence) Rui Figueira Sarah Rossie de Gasperis Sarah Vincent Serge Scory (RBINS) Sofie De Smedt (Meise Botanic Garden) Vince Smith, NHM, WP3 Leader Wesley Tack (Meise Botanic Garden) Elspeth Haston (Royal Botanic Garden Edinburgh

Notes

Milestone move updates - MS3,2 moved to April 2021 - discussion on scope and usefulness Overall an early Task in terms of delivery (all components due by the middle of 2021).

Summary of areas of exploration (see slides) *Helen Hardy*

DiSSCo needs across a range of organisations- three categories: leadership; digital, data & technology skills; specialisms (e.g HR, finance, legal etc)

Competency frameworks - prior attempts to work on collaborative competency frameworks (e.g. EUColComp) but not widely adopted. Reviewed other sectors for related roles - plenty for digital and tech roles. Others that need reviewing and integrating.

Commonality in competency frameworks - behaviours, specific knowledge and qualifications. Levels of competency and seniority/grades.

Ecosystem approaches to competencies - success profiles

Digital maturity in organisations - separate to individual competencies (**more work required**)

Case study - Luomus' digital transformation (see slides - eight key points) *Anne Koivunen*

Solid foundation with national and parent organisational support

Clear vision based on policy

Strong staff engagement - most challenging part, especially for staff who have been in roles for long periods of time

Objective-based approach

Reorganisation/restructure of resources

External funding including national GBIF node - FinBIF

Effective collections management system (internally developed)

Staff training

Lots of examples of good comms at conferences (Biodiversity Next) and other projects (ICEDIG)

Data, and finding skills/capacity - Helen Hardy

How do we find the skills we need?

Vocabulary for skills and expertise required

Task 3.1 Key discussion points - *Helen Hardy*

VS: CETAF discussion about EUColComp (**are there notes?**) was less suitable for a small/medium organisation.

AC: Proposal to use functional units - acknowledges hybrid roles profile in smaller organisations.

HH: How to apply? Still challenges with volume of data/competencies.

PM: In other projects (OpenUp! etc) did an exercise in defining roles - implementation struggled with having the same person in all roles. Issue with technical implementation of people in multiple roles (e.g. in PRINCE2 framework).

QG: Lack of involvement from HR/Directorate. Strategic approach required with decision making in specialisations for staff/entire organisations. How to address training if hiring is not possible?

HH: Link with WP2 and training

DK: **Evaluation framework in NL**. Broad context used for assessing salary scale across museum sector. Issue with people who do not fit neatly within a role (e.g. technical staff). [This example of a framework also shared by Elsa in chat: https://www.bls.gov/ooh/education-training-and-library/curators-museum-technicians-and-conservators.htm]

PG: Training gap identification. Sharing staff was discussed after EUColComp.

HH: Summary - plenty of examples, not many well-used (? At least in our sector) but good examples of how they are applied / can be beneficial if applied properly.

How to make tailored recommendations per subtask - **timing/resource requires a tool rather than manual one-by-one approach**. **Feedback on self-assessment**...

VS: Policy self assessment and SYNTHESYS CSAT (Collections Self Assessment Tool). What is useful?

EH: Tried digital culture compass. Found scope is too broad. Treats institutional collections as one - in reality they are managed separately. Need a tool that can be used at a subcollections basis (LL: for the archetypal collections like entomology, zoology, botany etc).

PM: CSAT has been useful and would be useful in smaller university collections. Can it be made available online again? Good to have an update.

DK: Provide **lightweight mechanism for personal self-evaluation** to help build personal profiles across individual institutions and wider DiSSCo Prepare. May help build a critical mass.

HH: Institutional self-assessment model may be useful and allow self-definitions of levels. Could ask whether certain competencies are reflected *somewhere* within the institute.

VS: Examples including Google Scholar Profile, LinkedIn, ORCID. No standard applied and unilaterally self-declared. **Could we propose our own metadata schema within these profiles**. This may provide a machine readable way to gather and amalgamate the data.

EF: **Teleworking suitability of jobs important**. WFH may necessitate different/new skills.

HH: Discussed ^^ in 3.3.

PA: Herbarium [which herbarium?] has one person dedicated to it - highschool education and in their 60s. General situation for many herbaria in Portugal. Self-assessment for collections with small staff numbers may be useful for knowledge/expertise sharing.

HH: Could look at LinkedIn data.

QG: Proposal by MeiseBG to do ^^

VS: Poll by VInce for indicating LinkedIn use (Y5, N1) + chat

POLL FOR LINKEDIN USAGE: https://strawpoll.com/jqceo724y
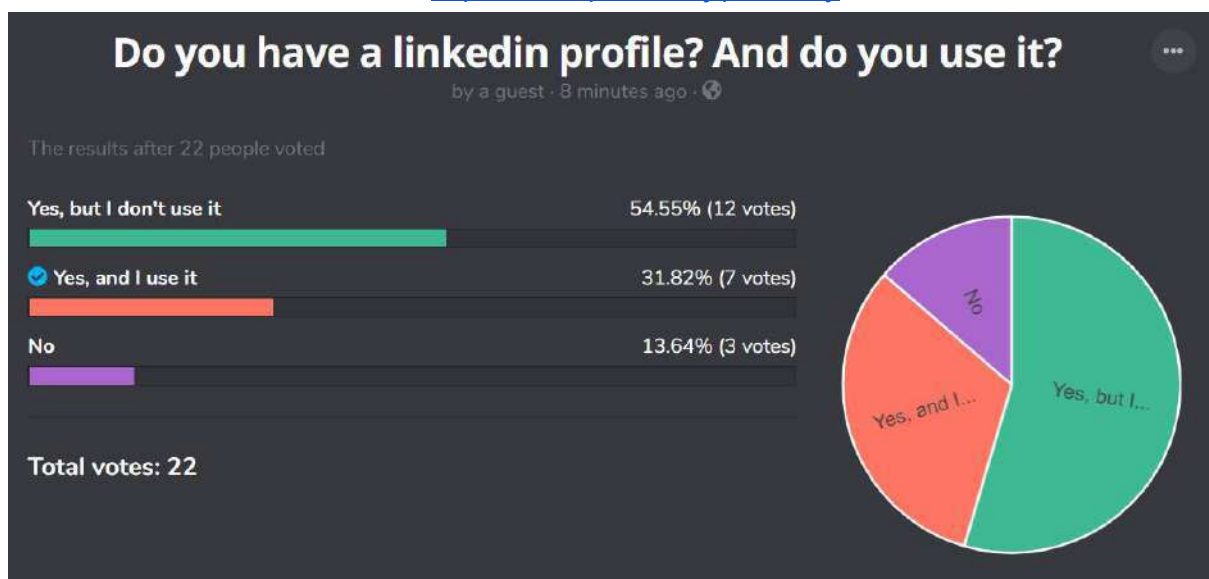
PM: National profile in use in BE - incentivisation to use. Check : https://researchportal.be/en

WA: People should be responsible for their own information - technically we can standardise and potentially use a DiSSCo extension on ORCID. Currently a **DiSSCo proposal to extend ORCID**. Shout out to **ELViS - European Loans and Visits System** elvis.dissco.eu

HH: Mindful of integration and collaboration with ELViS in Task 3.1

Task 3.3 - Discussion points - *Helen Hardy*
Distributed working
Secondment process
Examination of models to support secondment

HH: Only (successful) example in the UK culture sector known currently was philanthropically-funded scheme funded short-term secondments to V&A.
VS: Focus when writing tasks on (technical) development. Huge change is Covid and remote working - is the original vision still valid (physical movement of staff).
AH: **Substantial amount of work required for DiSSCO infrastructure. Likely to be fully occupied with existing day jobs. How can they be "released" from their day-to-day commitments. How will these scarce pools of people be harnessed for the benefit of DiSSCo**.
HH: Considering shared interest/insitutational interest.
QG: Have had subcontracting to allow for others to develop/work on a shared collections management system. Some experience. **Role for secondment outside of developers. External specialists good to have when implementing change management - provides neutral expertise**.
HH: Like a lighter touch approach to audits.
CW: How can we guarantee developers devote time to DiSSCo Project. Helpful to have a small number of meetings to assist with the team working for DiSSCO architecture.
LL: Propose call like MOBILISE STSM but inverted for specialists to visit.
HH: Possible internal consultancy type model?
PM: MOBILISE example - bird specialist wanted to apply to help a young new curator with their African bird collection . COST encourages mobility of younger professionals/researchers. So doing the opposite would need to be discussed with the COST office. There are also practical elements that it is easier to move the specialist to the collections, than the collections to him with the young researcher. (did not apply due to COVID so do not know if would have been accepted)
AK: Still important to consider the ways to improve online secondment. Easier to "lend" a specialist if there is not physical/full-time disruption. May not be so efficient but more balanced incentivised.
VS: **Need to understand and recognise incentivisation for secondment**.
DK: Financial incentives for institutes does not work (evidence?). **Provide a system of internal credits - allows you to "purchase" expertise from the consortium**.
HH: Distribution of financial incentives.
AH: Will come up in WP6 discussions.
AC: System not good for small organisations to get credits. Would not work evenly for small and big organisations.

Secondment examples : Marie Curie RISE (Staff exchange) + other Marie Curie programs
Daubenton project Life long learning for collection Managers
COST short time scientific missions
SYNTHESYS TA (more research than Capacity building, if training component too large, are not selected.
National funding mechanisms exist as well (bilateral agreements between countries
Also look at EURAXESS mobility opportunities

(works if external funding available)

Some on the job trainings available (ie Meise Seed bank management training which has an optional on the job training part)

Task 3.2 AHM Meeting Agenda & Notes

Background

The formal title for Task 3.2 is "Collate, refine and implement best practices for data mobilisation at the institutional level to develop the DiSSCo plan for data mobilisation and curation pipelines".

The full task description is provided in the linked planning document along with the subtask descriptions.

Participants

PLEASE ENTER YOUR NAME ON A NEW LINE BELOW

Laurence Livermore
Helen Hardy
Mil de Reus
Anton Güntsch
Esko Piirainen
Mathias Dillen
Patricia Mergen
Patrick Semal
Heli Fitzgerald
Ann Bogaerts
Anne Koivunen
Wouter Addink
Ana Casino

Alex Hardisty
Elspeth Haston (had to leave at approx 12.50 GMT)
Josh Humphries
Maarten Trekels
Maria Joao Santos
Quentin Groom
Vince Smith
Serge Scory
Sarah Rossi de Gasperis
Wesley Tack
Judite Alves

Ville-Matti Riihikoski
Frederik Berger
Sofie De Smedt
Tina Loo
Pedro Arsenio
Roger Hyam
Tania Walisch
Elsa Fontainha
Kari Lahti
RosaRosa?
Rob Cubey
Carole Paleco (RBINS)
Tania Walisch (MNHNL)

Draft Agenda

- Task Overview
- Discussion
    - Current state of institutional digitisation
    - General state of natural science collections digitisation
    - Task logistics

Minutes

Introductions

[add link to Laurence Livermore slides for introduction]

Questions and discussion

Questions shared in chat:
What are you currently digitising (or planning to do when you return on-site)?
Has COVID affected your institution's plan or outlook on digitisation?
How is your institution monitoring digitisation at the moment?
What is everyone digitising (or planning to digitise) beyond the well-tested and fairly well-published workflows (pinned insects, microscope slides, herbarium sheets)?
Do you have pre-digitisation protocols? How do you assess and prioritise collections?
Do you have well-documented internal protocols for your workflows? How much of them can be generalised? Do you have documented hardware and software? How much post-processing and temporary (logistics/operational) metadata capture do you use?
What are your pain points/constraints in digitising data? What do you feel could most easily be improved (considering costs and benefits)
How do you publish and share digitised data?

What is the current state of institutional digitisation?
LL asked for a quick summary of key projects/programmes now (or on return to Museums depending on Covid restrictions).

Luomus not too affected by Covid restrictions.
A lot of work to digitise insect specimens.
Use of conveyor belts for imaging.
Also transcribing.
Plan to acquire a CT scanner this year, and two more imaging stations for lichens and mosses - photographs plus OCR to read labels (new for Luomus - this collection has more typed labels so should be more straightforward).

Meise in 2nd mass digitisation project, to digitise 1.2m specimens which are the general herbarium using Picturae. Had to stop last March but then restarted after a month with Covid precautions such as barriers - brought rate down from 4,000 per day to 3,000. Imaging now expected to continue until May. Transcription keeps proceeding using the team in Surinam. In-house digitisation of micro-algae also stopped but now restarting - only allowed to access the collection one day a week at present. Transcription did continue at home, using citizen science platform from first digitisation project on Belgian collections. As a citizen Science project, this saw high uptake while people were at home during Coronavirus.

Berlin currently not able to digitise on site so focusing on transcription. Running a Zooniverse project on bee labels (a large one). When back on site will start up in entomology - tendering

for a system with throughput of 5,000 specimens per day similar to a herbarium approach.
Also work on dry invertebrates (mollusca and fossil inverts) - have a new camera system, not yet measured but expected to speed up in house digitisation of this material.
Also developing 3D and High res imaging.
CT scanner installed and expected to be operational next month.

RBG Edinburgh most digitisation in house (core funded). Recently worked with a local company to replace light boxes. Also purchased new cameras to go with these.
A major migration to Specify from old CMS - working to develop data entry tool for this for digitisation.
Crowdsourcing projects with DigiVol and also working with Zooniverse
CMS migration slowed work as much or more than lockdown.
Work on MIDS(minimal information about digital specimens standard) with CETAF et al - keeping it simple on imaging (is there any image or not?).

This task will need to work with the CETAF Digitisation Working Group. A key connection.

Lisbon - research infrastructure in Portugal for biodiversity data. Some extra funding in last 3 years for digitisation in collections, including herbarium, zoology and a little in geological collections. A lot still to be done in standardising workflows - different collections do own programmes at present. Funds running down and not sure how will fund digitisation in future (affects different departments/institutions in Lisbon) - makes it hard to keep any kind of 'mass digitisation' rhythm.

NHM mostly remote working last spring-summer and now, focusing on transcription. When back on site will be focusing on Synthesys+ Virtual Access projects - freshwater flies and bats. First zoological project with the bats. Also looking at workflows for carpological herbarium collections; and maybe fish depending on staff numbers.

RBINS: Micro CT is running normally because it is a separated building with possible social distancing. The digitization on site is allowed 1 day a week with homework for the treatment of the data using remote applications like "teamviewer". The manual data and metadata encoding continue normally but as homework with quality control of the encoded data. Adaptation of the Collection Management system to CETAF Stable identifier and IIIF compliance for the image server (common work with Africa Museum Tervuren)

Is anyone mass digitising anything other than slides, herbarium sheets and insects who hasn't already spoken?
NB there isn't a convenient definition of mass digitisation (although there is one from ICEDIG - shared in meeting chat by Alex Hardisty). Just looking at cost or throughput can be misleading but is a broad guide.
Noone raised anything they felt would be considered mass digitisation not already mentioned.

NHM mentioned constraints on getting data into CMS at mass scale. Question to the group for other blockers/pain points?

Berlin - hard to find funding to scale up or to hire people on short term contracts with relatively low sector pay. Work with industry also challenging.

Luomus - In a relatively fortunate position, but roles can be challenging e.g. where staff have other responsibilities as well as digitisation. Can take time to resolve any hardware or software issues.

RBGE - Issues with importing data into CMS, particularly in the context of migration. Relates to the nomenclatural backbone (would like to import International Plant Names Index). Have electronic data on many historical collections - couldn't previously import this but that may become easier with the new system. Similarly for crowdsourced data.

Pat Mergen - **when DiSSCo is a legal entity would it be helpful to issue tenders above institutional level? Could support outsourcing. A question for later date.**
Still requires funding of course. Examples do exist of individual and small groups of organisations tendering.
(https://ec.europa.eu/digital-single-market/en/pre-commercial-procurement) linked to Task 4.4 of DiSSCo Prepare

Has anyone used or adapted a documented workflow (e.g. from a paper)?
How widely are barcodes used and should that be assumed best practice for DiSSCo?
WA - yes for newly generated digital material, would advocate adding barcodes as part of digitisation process.
Meise - found papers useful, particularly on imaging quality and data standards etc. Helped in preparing tender documentation. Also looking at NHM microscope slide workflow. Standards includes broader ones than our community e.g. TIF standard or colour quality control etc. Sector-specific standards and points tend to be more known or accessible already? Could use this task to complete write-up and publication of some workflows from ICEDIG that were not completed.
There was an ICEDIG deliverable on quality control for imaging.
LInks to some material from prior projects are in the planning document.

This task will need to work with the knowledge base. Need to ensure a consolidated collection within that e.g. for SoPs. Publication is a bit separate.
Knowledgebase will have some metadata and API.
IDigBio also have resources - how could we help to bring resources like this together and make them more searchable? And have data e.g. on when procedures were last updated or similar.
ACTION - review key IDigBio resources as part of this task and provide a way to access the most useful?
HH - liaise with IDigBio on this? They have a very wide base of useful material e.g. they have some relevant to WP3 as well]
ACTION - Get in touch with WP5 on how Knowledgebase can be used consistently by this WP

WA - where is the crossover from having best practices to more action to encourage their use?
Probably not in scope for this task - a recommendation - but could be part of work e.g. in MIDS? Could look to set more requirements? Perhaps this goes beyond the Prepare phase of the project? A transition towards the end?

What about recommendations on the process for conversion of best practices etc to use? (even if the actual conversion is premature)
A lot depends on the granularity of SoPs - often reasons to deviate from a detailed process. Also a constant change - need to reflect the need to accommodate innovation in a sustainable way. [HH note - does the IDigBio wiki do this at all?]

Task logistics

How frequently should the group meet? Fortnightly agreed. **ACTION LL to schedule**

Cetaf Digitsation Working Group and ISTC will be very important.

How open is the MIDS discussion? Takes place in TDWG task group chaired by Alex Hardisty and Elspeth Haston - can join the email group on request. Also the CETAF DWG who feed in best practice suggestions etc. Open to CETAF members.
There is a Github repository here: https://github.com/tdwg/mids

Alliance for Biodiversity Knowledge and others consulting from Feb about digital specimens and extended specimens - one of 5 strands of this will be about mobilising FAIR specimen data. Alex H co-moderating that as is Wouter Addink and Barb Tiers (New York).

Alex Hardisty chat comments:
Alliance For Bio blog post on forthcoming global consultation:
https://www.allianceforbio.org/post/converging-digital-and-extended-specimens-towards-a-global-specification-for-data-integration .
Background document on the consultation:
https://docs.google.com/document/d/1mtjLD7Zpf73apLajW8qStA3U0Gpz2ZGazZkjN2Vp3Ew/edit
Strand 1 in the consultation, on "digitizing/mobilizing FAIR data for specimens" is relevant for task 3.2 participants. Contributions would be welcome. Consultation is likely to be open from 16th February to 5th March.
TDWG Task Group on MIDS: https://github.com/tdwg/mids and the current draft of the specification:
https://github.com/tdwg/mids/blob/working-draft/current-draft/MIDS-definition-v0.12-03Nov2020.md . TG MIDS aims to meet monthly a few days after the CETAF Digitisation WG meeting, which also aims to meet monthly. https://cetafdigitization.biowikifarm.net/cdig/

Luomus have volunteered to work on standardised ETL, Meise on pre-digitisation curation, and NHM on SoPs.
Who could lead subtask on digitisation monitoring? Edinburgh willing.

NHM will review existing material from IDigBio and consider gap analysis across this, ICEDIG etc on digitisation workflows/SoPs. Areas where workflows don't yet exist or have a lot of variance. Probably also extends to quality assurance.

Luomus - would like to understand the subtask goals more clearly around ETL and start with a discussion about this.

CMS import as discussed above is of interest but also other aspects.

Esko - The overall architecture is relevant e.g. links to other systems/platforms etc

Some links to other work packages. WA - some work on CMS system requirements Task 6.1 led by MfN and WP8 task on OCR.

ACTION to coordinate this sub task with these or other relevant tasks.

Need a clear separation in steps of digitisation process - e.g. so that OCR could take place in a distinct tech 'pipeline'?

Pre-digitisation and monitoring similarly will need a gap analysis / analysis of previous work, and discussion with institutions.

Check SYNTHESYS Plus task 2.1 on policies, which addresses also needs in terms of Pre and post digitization notably in the framework of Virtual Access and future Digitization on Demand within DiSSCo.

Anyone who is closely involved in digitisation or in technical aspects of data mobilisation please do make themselves known.

Work to create a list of principle contacts for key topics is underway (by CSO? WA aware of this)

JA - There is a helpdesk aspect around what should be provided by the helpdesk - e.g. could include documentation?

Would this include the handbook? Not started yet but likely.

All zoom meeting chat for reference:

From Anne Koivunen to Everyone: 12:06 PM
Luomus is also a partner
From mariajuditealves to Everyone: 12:06 PM
Lisboa is present. But it is Rui Figueira that is supposed to participate in the future
From Pedro Arsénio to Everyone: 12:08 PM
Hi everyone, I'm also from Lisboa (Pedro Arsénio).
From Vince Smith, NHM to Everyone: 12:09 PM
Thanks Anne and Judite and Pedro
From Laurence Livermore to Everyone: 12:13 PM
What are you currently digitising (or planning to do when you return on-site)?
Has COVID affected your institution's plan or outlook on digitisation?
How is your institution monitoring digitisation at the moment?
What is everyone digitising (or planning to digitise) beyond the well-tested and fairly well-published workflows (pinned insects, microscope slides, herbarium sheets)?
Do you have pre-digitisation protocols? How do you assess and prioritise collections?
Do you have well-documented internal protocols for your workflows? How much of them can be generalised? Do you have documented hardware and software? How much post-processing and temporary (logistics/operational) metadata capture do you use?

What are your pain points/constraints in digitising data? What do you feel could most easily be improved (considering costs and benefits)

How do you publish and share digitised data?

https://docs.google.com/document/d/1EWyEHNNdm6UoL0gnffmCH2pliTE0HH0kJteJeet6C1M

From Mathias Dillen to Everyone: 12:20 PM

https://www.doedat.be

From Wouter Addink to Everyone: 12:25 PM

For the insect digitization in MfN, are any of the new technologies being used as researched in ICEDIG?

From Frederik Berger (MfN) to Everyone: 12:28 PM

@Wouter, yes, we prepared the tender based on the recommendations of the ICEDIG output. We expect some new technological developments from the tender, which would also include focus stacking in mass digitization.

From Mathias Dillen to Everyone: 12:31 PM

You can't raise your hand if your host

you're

From Alex Hardisty to Everyone: 12:36 PM

In the ICEDIG project we defined mass digitization as follows: Mass digitization: An activity where entire collections, or their distinct major parts are digitized from one end to the other, without selecting individual specimens. Mass digitization is characterised by technological and procedural frameworks based on automation (hardware and software) and enrichment (link-building), with workflows at industrial scale, i.e., processing millions of objects at low cost.

From Laurence Livermore to Everyone: 12:36 PM

Thanks Alex :)

From Patricia Mergen to Everyone: 12:44 PM

https://ec.europa.eu/digital-single-market/en/pre-commercial-procurement

From Esko Piirainen to Everyone: 12:44 PM

From technical point of view, not neccesarily related to (MASSdigitazion in particular) sharing large-scale original images is a pain-point: no way for us to do this now except extracting individual images manually

From Laurence Livermore to Everyone: 12:46 PM

General topics
* What is the current state of information sharing for processes and workflows?
* Where should we publish workflows? What would be a sensible and sustainable way of doing so?
* How is the wider community converging on barcode usage?
* How can we support digitisation workflows and take into account a range of conditions (e.g. small, institutions lacking IT support, staff shortage, limited resources for equipment, etc)?

From mariajuditealves to Everyone: 12:46 PM

Thanks Alex

From Quentin Groom (Meise) to Everyone: 12:47 PM

We used many papers on image quality and digital standards before setting up our workflows.

From Elspeth Haston to Everyone: 12:48 PM

I'm sorry, I need to leave. Apologies for not being able to stay right now and contribute. I'm happy to meet regularly on this Task - and happy to help the communication between CETAF Digitisation Working Group and MOBILISE Working Group 2.

From Laurence Livermore to Everyone: 12:52 PM

https://docs.google.com/document/d/12ywJGyo2_4ps8ogBwJ0HP-arVlWld-BpZ8tmZMd1LvA

From Quentin Groom (Meise) to Everyone: 12:53 PM

Nieva de la Hidalga A, Rosin PL, Sun X, Bogaerts A, De Meeter N, De Smedt S, Strack van Schijndel M, Van Wambeke P, Groom Q (2020) Designing an Herbarium Digitisation Workflow with Built-In Image Quality Management. Biodiversity Data Journal 8: e47051. https://doi.org/10.3897/BDJ.8.e47051

From Alex Hardisty to Everyone: 12:55 PM

D3.1 Quality management methodologies for digitisation operations, April 2019. https://doi.org/10.5281/zenodo.3469521.

From Vince Smith, NHM to Everyone: 12:58 PM

Knowledgeable session is Tuesday 15:00

From Wouter Addink to Everyone: 12:59 PM

Knowledgebase session tomorrow 15:00CET

From Laurence Livermore to Everyone: 01:00 PM

Task Logistics

* What is our timeline and agreed actions?
* How can we collaborate with key groups like the CETAF DWG?
* How frequently should we meet to discuss and share progress.?
* What form will the output(s) take?
* How can we make them as useful and sustainable as possible?
* Decide upon a subtask 3.2.4 "Digitisation Monitoring" leader

From Ana Casino to Everyone: 01:07 PM

We could help in that regards, to come back to community and making this as a recommendation

CETAF WGs may help a lot in this regard

Digitization WG coordinator is Elspeth Haston. MIDS is at the core of its work

From Josh Humphries to Everyone: 01:08 PM

here's the MIDS github page: https://github.com/tdwg/mids

From Alex Hardisty to Everyone: 01:11 PM

Alliance For Bio blog post on forthcoming global consultation: https://www.allianceforbio.org/post/converging-digital-and-extended-specimens-towards-a-global-specification-for-data-integration. Background document on the consultation: https://docs.google.com/document/d/1mtjLD7Zpf73apLajW8qStA3U0Gpz2ZGazZkjN2Vp3Ew/edit

Strand 1 in the consultation, on "digitizing/mobilizing FAIR data for specimens" is relevant for task 3.2 participants. Contributions would be welcome. Consultation is likely to be open from 16th February to 5th March.

TDWG Task Group on MIDS: https://github.com/tdwg/mids and the current draft of the specification: https://github.com/tdwg/mids/blob/working-draft/current-draft/MIDS-definition-v0.12-03Nov2020.md. TG MIDS aims to meet monthly a few days after the CETAF Digitisation WG meeting, which also aims to meet monthly. https://cetafdigitization.biowikifarm.net/cdig/

From Patricia Mergen to Everyone: 01:23 PM

In SYNTHESYS 2.1 (on policies) they have a part on needs of pre-digitization
From sofie (Meise Botanic Garden) to Everyone: 01:24 PM
thanks Patricia
From Ana Casino to Everyone: 01:27 PM
Good point Judite
From sofie (Meise Botanic Garden) to Everyone: 01:27 PM
WP 3.2 is also related to T1.3

ALL HANDS MEETING – T4.1
DiSSCo COST BOOK - METHODOLOGY ASSESSMENT
OUTCOMES & NEXT STEPS

January 22nd 2021

AHM - T4.1 - OUTCOMES FROM DISCUSSION

The experimental method proposed by MNHN is **complex and does not reach a consensus** among DiSSCo members.

Many questions are still pending:

- How to **simplify** the method?

- How to consider the **diversity of institutions (administrative, accountancy, size)** that constitute the DiSSCo landscape?

- Do we **integrate living collections** into DiSSCo perimeter?

- How to **formulate a flat rate** to assess indirect costs that corresponds to each institution?

- Is this **method worthwhile** in terms of cost-benefit?

- Is it **possible to cover all indirect costs**, knowing that it represents a large range of costs and numbers (m2, staff time, etc.) that are not easy to gather?

AHM - T4.1 - NEXT STEPS

1. A **method will be developed** within MNHN to assess indirect costs.

2. This method will be **applied to DiSSCo member institutions**.

3. With these results, a **ratio may be developed**.

4. This **ratio could vary** depending on the type of institution (Botanical garden, Museum, size, country, etc.).

"Bringing the irreplaceable data stored in natural science

collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

The preparatory phase project of DiSSCo
Research Infrastructure - Distributed
System of Scientific Collections

https://www.dissco.eu/prepare/

ALL HANDS MEETING

WP4 (Business framework) – T4.1 (DiSSCo Cost Book)

METHODOLOGY ASSESSMENT – INDIRECT COSTS

MEETING MINUTES - Thursday, January 21st 2021

Speaker: Michel GUIRAUD (MNHN)

Participants (incomplete):

MNHN : Michel GUIRAUD, François DUSOULIER, Eva PEREZ, Salomé LANDEL

CETAF : Ana CASINO, Marie-Laure KAMATALI, Celine CASSARINO

DiSSCo CSO: Tina LOO, Dimitris KOUREAS, Eva ALONSO

Meise : Patricia MERGEN, Sofie DE SMEDT, Mathias DILLEN, Wesley TACK, Quentin GROOM

SGN: Hanieeh SAEEDI, Peter WARTH

BGBM: Eva HAFFNER, Agnès KIRCHHOFF

UniFI: Luca BARTOLOZZI, Rosarosa MANCA, Lorenso CECCHI

NHM: Laurence LIVERMORE

NM: Jiri Frank

TIME	TOPIC
9:00 – 9:10	Presentation of WP4.1 (DiSSCo Cost Book)
9:10 – 9:30	Introduction to indirect costs
9:30 – 10:00	Discussion on indirect costs
10:00 – 10:30	Methodology for the calculation of DiSSCo's indirect costs

UPORTO: Maria Joao FONSECA

AGENDA

I. MAIN DISCUSSION, HIGHLIGHTS AND DECISIONS
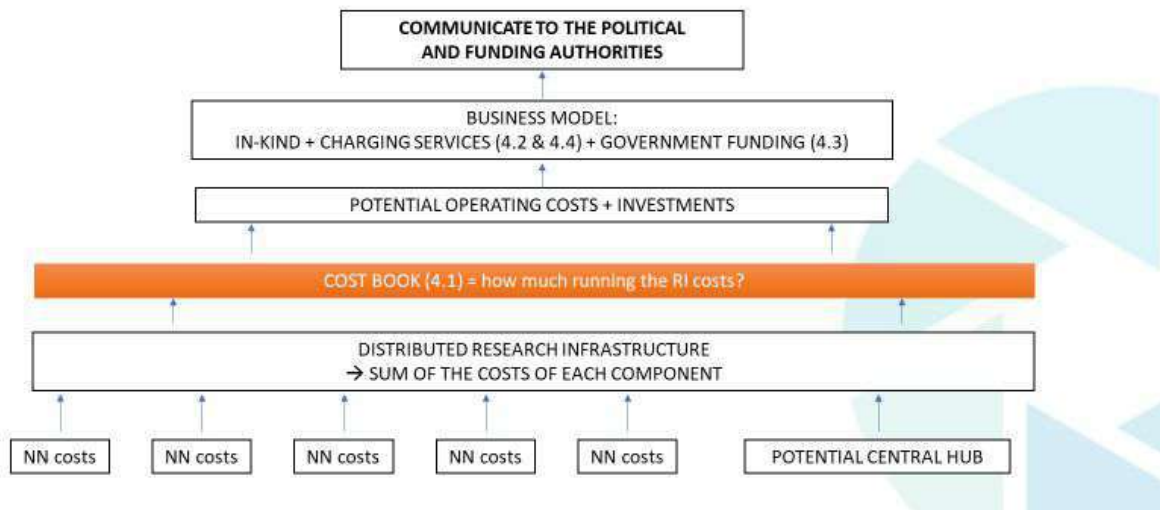
Michel Guiraud presented the objective of the cost book: it aims at assessing DiSSCo costs to produce a narrative for political and funding authorities. As DiSSCo is a distributed infrastructure, costs are not centralized within DiSSCo Hub because its services depend most of the time on local institutions that hold the collections. The cost book should gather costs from national nodes.

In the first place, T4.1 had to delimit **DiSSCo Perimeter**. In September 2020, nine areas of costs have been presented to the NNs (DPP_MS4.5). These areas go from IT infrastructure to governance. They are meant to cover all DiSSCo activities. These activities represent the basis of DiSSCo business model. The idea is that the RI not only depends on government funding but also provides services to external users as well. Such services shall be developed under T4.2 (cost model for charging services).

With such objectives, T4.1 aims at providing a tool, simple to use, to report and pilot the costs of the RI. It would be a methodology to assess all costs that are part of DiSSCo perimeter. Two main types of costs have been identified: the direct and indirect costs.

DIFFERENCE BETWEEN DIRECT COSTS AND INDIRECT COSTS
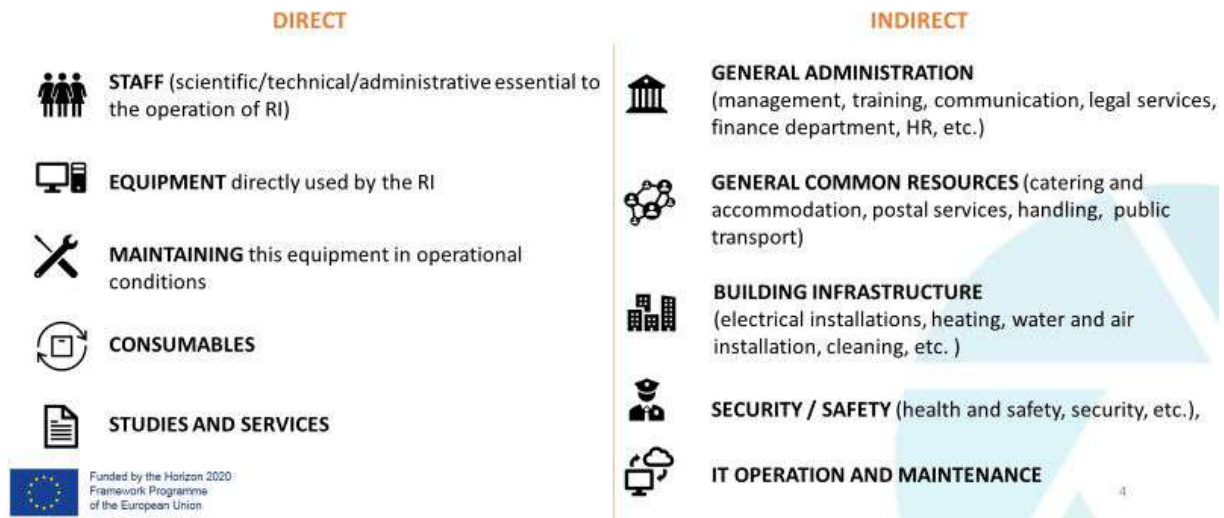
Direct costs are the one directly connected to the project. For DiSSCo, it relies to IT infrastructure, IT tools, data curation, digitisation, physical access to collections, analysis, conservation costs, physical access to collections and governance.

It includes staff time (scientific / technical administrative) essential to the operation of the RI. For instance, with DiSSCo, this includes collection managers, IT developers, etc. Other direct costs are equipment and their maintenance, consumables and studies to improve the operation of the RI.

Indirect costs are costs that surround direct costs. It gathers general administration, general common resources, building infrastructure, security, IT operation and maintenance. Such costs provide the environment that permit the work done for the infrastructure. For instance, staff requires HR to follow their work, and electricity to make the equipment work.

There is a **pending question regarding IT operation**, whether it is fully or partly connected to DiSSCo and whether it is a direct on an indirect cost.

WHAT ARE INDIRECT COSTS?
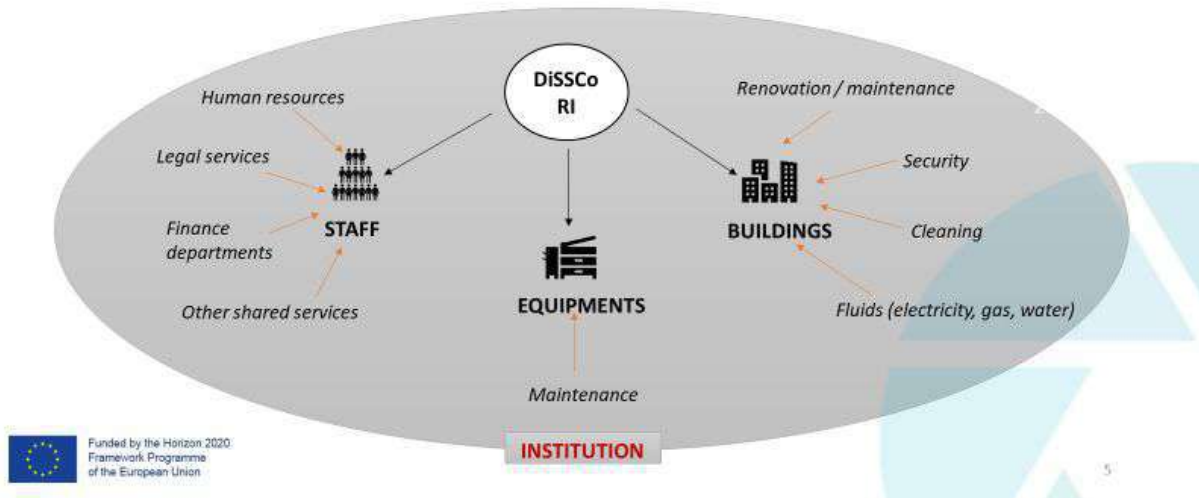
In the context of DiSSCo, on the one hand, indirect costs are services that surround the staff hired for DiSSCo. On the other hand, there are the costs to maintain building (renovation, security, housekeeping), fluids. These indirect costs are related to surfaces.

HOW CAN WE ESTIMATE THOSE INDIRECT COSTS?

€ ASSESS THESE COSTS WITHIN THE INSTITUTION
How much HR services cost per year? (etc.)

DEVELOP A RATIO: COST PER STAFF MEMBER OR PER SQUARE METRE

= GROSS ANNUAL SALARY CHARGED OF THE STAFF INVOLVED IN THE PROJECT * RATIO
or / and
= SQUARE METRES OCCUPIED BY THE PROJECT * RATIO

Within MNHN, a pilot project has been conducted between October and December 2020. **Two types of indirect costs have been identified:**

- **The one that surrounds the staff**: it covers administrative work. It may be calculated by dividing the overall cost of administration with the number of employees. It gives an average cost per staff member.
- **The one that surrounds the buildings**: it covers building maintenance, electricity, etc. It may be calculated by dividing the overall cost of building maintenance with the number of square metres of the institution. It gives an average cost per m².

Based on the 1st discussion that we had with NNs, not everybody has an analytical accounting to provide the exact cost of fluids. Within MNHN, people are working on that. Michel Guiraud explained that it might be possible to have some numbers, but certainly not for everything.

To prepare the AHM session, **T4.1 asked to a panel of institutions across the European union, 4 figures:**

We selected 2019 because it is the last closed accounting year.

1. Overall annual budget for the administration (HR, finance department, legal service, etc.) in 2019?
2. Annual cost of maintenance (building repairs, servicing) and fluids (water, gas, electricity) in 2019?
3. How many people were working in the institution in 2019?
4. What is the total surface area of the institution (in square metres)?

These numbers were supposed to be easy to find. They can be asked to the head of the finance department, or found in the annual report.

In one week, T4.1 managed to gather the following numbers. There are differences between ISBAS, NHM, MNHN and BGBM. With these results, Michel Guiraud explained that he finds them interesting because the difference between the numbers provided are from one to ten, and it is "basically not much". Concerning the difference in purchasing power between European countries, he proposes the big mac index. If this index is used, the difference is from one to six. Michel Guiraud explained that these results demonstrate that this approach is the right one.

	NHM	ISBAS	MNHN	UniFI	BGBM
ADMINISTRATION	9 796 640,00 €	162 360,00 €	15 598 155,00 €	No info	639 395,00 €
BUILDING	14 148 958,88 €	148 600,00 €	7 654 703,00 €	No info	3 787 062,00 €
STAFF	936	136	2685	45	213
M2 (without gardens)	127 601 m2	4 912 m2	129 000 m2	13 800 m2	12 477 m2
SUPPORT COST / PERSON / YEAR	10 466 €	1 193 €	5 809 €	No info	3 001 €
SUPPORT COST / M2 /YEAR	110,88 €	30,25 €	59,34 €	No info	303,51 €

At this stage, the results from DiSSCo partners can be classified as followed:

	NHM	ISBAS	MNHN	UNIFI	BGBM	SGN	RBINS
1	exact value	Good estimation (uncertainty is 10%)	Approximation (say uncertainty is 20-50%)	No info	exact value	No info	No info
2	exact value	Good estimation (uncertainty is 10%)	Approximation (say uncertainty is 20-50%)	No info	exact value	No info	No info
3	exact value	Good estimation (uncertainty is 10%)	Approximation (say uncertainty is 20-50%)	Good estimation (uncertainty is 10%)	exact value	No info	No info
4	Good estimation	Good estimation (uncertainty is 10%)	Approximation (say uncertainty is 20-50%)	Good estimation (uncertainty is 10%)	exact value	No info	No info

This approach is a first step, and the cost book is built step by step. Michel Guiraud also recognized that people involve in DiSSCo are most of the time scientists and not accountants. Meanwhile, the working plan for T4.1 is to assess DiSSCo indirect and direct costs. It relies on the participation of each nodes: at some stage, T4.1 will need their numbers. T4.1 is supposed to create a methodology that ease the exercise.

During the session, the main discussion was on the feasibility of this method. Luca (Unifi) pointed out that his Museum belongs to the University of Florence. **The costs are calculated by the Uni and condensed with other costs that are not DiSSCo related.** He said that it is possible to calculate a proportion that will not be an exact value. They have also different types of buildings that are not directly connected to natural history collections. Luca also asked why salary are indirect costs. He added that indirect costs will not be calculated easily.

MG (FR): replied that staff is direct and indirect cost. Some people are directly connected to the project. Some others (housekeeping, HR) and indirectly connected to DiSSCo.

Patricia (Meise) noted that on the powerpoint square metres exclude gardens. She pointed out that **gardens represent a large part of the collections in Meise**. She added that for some institutions, parks are only ornamental, and for some others, gardens are part of the collection. The question of how to integrate living collections in DiSSCo is important. In Flander, at local level, living collections are included.

Patricia also presented the way EOSC cost model analysis tackles the question of indirect costs. It distinguished two ways:

- **Project approach**: indirect costs are often well understood and decided by the funder. There is an agreement between the local project runner and the administration. Different models exist:
  - the administration takes what corresponds to indirect costs, whatever the funder is saying and the project manager use direct costs and never see indirect costs.
  - the institution has its own ratio (10%, 20%) and whatever the funder says, even if they give another ratio, the institution takes what has been decided within the institution.
  - Other cases: institution does not take indirect costs at all and let the project manager use them as they want for the project

- **Institutional approach**: it's just staff whatever they do, they have a budget line for the personnel only = no distinction between direct and indirect costs = depends if you have an analytical accounting system

EOSC did not manage to have one approach on how much it will cost for its implementation = everybody gives their figures that are put in common. The way these numbers have been produced is heterogeneous.

Ana (CETAF): **a fixed ratio of indirect costs among DiSSCo institutions also depends on the legal model that will be chosen.** If it is an AISBL, it remains at the institutional level. If it is an ERIC, it belongs to the national level and the data shall be aggregated.

Patricia (Meise): with EOSC, they interviewed big aggregators about services they offer. They also interviewed institutions. The partners remain the institutions. For the Cost Book, an agreement has been found between Belgium institutions but no common calculation for indirect costs. The size and the type of institution are too different.

Eva (BGBM**): the level of complexity is high and it is difficult to standardized such costs among DiSSCo institutions.** All institutions might have different specialities. Important to talk about it. She added that it does not only deal with m² but also the staff that maintain the garden.

Patricia (Meise): adds that regarding living collections some costs will vary if it's indoor or outdoor. She underlined that heating is massive for greenhouses.

As part of the Synthesys project, they had to take 2 years of maintenance expenses, and decide what is link to access to collections. Patricia did it and it was a massive work. Not sure that all institutions are ready for this effort.

Dimitris (DiSSCo CSO): this morning presentation demonstrates that this exercise is very complex and maybe impossible. **At the European union level, indirect costs are not asked to be calculated. The easiest way would be to agree on a direct calculation of employment per unit (salaries, securities, pensions, etc.).** On top of that, apply a flat rate of indirect costs. It would apply to different institution type, it might lead to a very difficult conversation about how to define the different types of institution. **Another solution: apply a DiSSCo wide flat rate that would be easier and quicker to apply.** That ratio could be reviewed in the future, as a flat rate percentage. This could be good compromise and a much easier way to take indirect costs into account.

Ana (CETAF): **T4.1 should list the advantages and the disadvantages to assess indirect costs or to use a flat rate.**

Eva (BGBM): What kind of data will be asked from T4.1 to DiSSCo member institutions?

If we choose an average between all indirect costs from DiSSCo institutions, how fair would it be for institution that are above this average?

II.   NEXT STEPS

A method will be developed within MNHN to assess indirect costs. This method will be applied to DiSSCo member institutions. With these results, a ratio may be developed. This ratio could vary depending on the type of institution (Botanical garden, Museum, size, country, etc.).

III. CONCLUSIONS

With this meeting and the discussion that took place, we can conclude that the experimental method proposed by MNHN is complex and does not reach a consensus among DiSSCo partners. Some questions are still pending:

- How to simplify the method?

- How to consider the diversity of institutions (administrative, accountancy, size) that constitute the DiSSCo landscape?

- Do we integrate living collections into DiSSCo perimeter?

- How to formulate a flat rate to assess indirect costs that corresponds to each institution?

- Is this method worthwhile in terms of cost-benefit?

- Is it possible to cover all indirect costs, knowing that it represents a large range of costs and numbers (m2, staff time, etc.) that are not easy to gather?

A common standard across DiSSCo shall be found. It should be as close to the reality as possible but also not too complex to find. At the end, different ratios according to the type of institution could be developed. For example, a ratio for botanical gardens and a ratio for museums could be developed.

IV. REFERENCE DOCUMENTATION

DiSSCo PREPARE

- DPP_WP4_Ms4.5_Set_of_indicators_forming_the_basis_of_the_cost_book_September_24th_2020_1
- DPP_WP4_Ms4.5_Set_of_indicators_forming_the_basis_of_the_cost_book_September_24th_2020_2

ICEDIG

- D4.5 - Cost analysis of transcription methods - DOI: https://doi.org/10.3897/rio.6.e56211
- D8.1 - Conceptual design blueprint for the DiSSCo RI - DOI: https://doi.org/10.3897/rio.6.e54280
- D8.2 - Cost Book of the digitization infrastructure - DOI: https://doi.org/10.3897/rio

V. AHM_CHAT_WP4_21-01-21

De Patrick Semal à tout le monde: 09:16 AM

In some Institutions like RBINS, we already have a cost model based on the average cost for a staff profile and a fixed ratio of this salary for the indirect costs includin or not the use of the collections. This cos model is used for the services activities of the RBINS. I suppose that many instituytions have already this information.

It could be easier to use just this cost model than to redo several times the same work ...

De Ana Casino à tout le monde: 09:29 AM

Generally Indirect costs cannot be assigned, applied, related to one specific task or outcome, but on the contrary need to be distributed among the bunch of those activities or outcomes those costs contribute to/ support

Terminology needs to be explained. In case of the buildings, the cost ...does it a net cost after mortgages or other? the exclusion of garden is linked to the fact that mortgages apply only to the buildings not to the ground.

De Luca Bartolozzi (Florence NH Museum) à tout le monde: 09:32 AM

we also have in Florence 23,800 square meters of the Botanical Garden which were not included in the table

De Ana Casino à tout le monde: 09:33 AM

In order to collate homogeneous figures across the partners

De Francois DUSOULIER (MNHN) à tout le monde: 09:35 AM

@Ana you are right, terminology is a work in process for the report

De Luca Bartolozzi (Florence NH Museum) à tout le monde: 09:36 AM

our University calculates indirect costs on a project at 25%

De a.kirchhoff à tout le monde: 09:44 AM

our University calculates indirect costs on a project at 25%
but we have a certain calculation model to calculate full costs for research services

De a.kirchhoff à tout le monde: 09:50 AM

this model is based on costs per staff hour

De Ana Casino à tout le monde: 09:55 AM

+1 Michel, what is in DiSSCo is an essential question,

living collections, exhibitions...are those under a digital driven RI? those are questions to address

De Luca Bartolozzi (Florence NH Museum) à tout le monde: 09:58 AM

we can do e.g. genetic research on living collections… why not to consider them in DiSSCo? What is the opinion of the Botanical Gardens that are partenrs of DiSSCo?

De l.cecchi@unifi.it à tout le monde: 10:00 AM

I think that living collections must be definitely included, they are as much "instable" as… Persons!

T4.2 Next Steps

The key actions/next steps are as follows:

1. NHMUK - Review all dependencies, agree subtasks between partners, then finalise with WP4 lead and Eva. (Feb 2021)
2. NHMUK/MeiseBG - Discuss scope of planned work and synergies with other tasks/subtasks (Feb 2021)
3. MHNH/MeiseBG - Write and review a common document for technical subcontracting work - NHMUK to review and add T4.2 requirements (Feb 2021)
4. NHMUK/DiSSCoTech/MNHN - Meet to discuss technical platform requirements and WP6 [Claus Weiland + Sharif Islam] (Feb 2021)

Task 4.2 AHM Meeting Agenda & Notes

Date: 2021-01-21

Background

See linked [Milestone discussion document](#).

Participants

31 total
PLEASE ENTER YOUR NAME ON A NEW LINE BELOW
Laurence Livermore (NHMUK)
François Dusoulier (MNHN)
Lorenzo Cecchi (Natural History Museum, Florence, Italy)
Claus Weiland, SGN
Sofie De Smedt (Meise Botanic Garden)
Salomé Landel (MNHN)
Eva Alonso (Naturalis)

Notetaker(s):

Elizabeth Devenish

Draft Agenda

- Task overview
- Discuss task scope & relationship with other tasks/projects
- Discuss subtasks & subtask leadership
- Agree next steps

Minutes

Meeting started: 12:00pm

1) Introductions - Laurence Livermore

2) Task Summary - Laurence Livermore
   - 11 person months split across partners.
   - Who will build and maintain data storage

3) Task Timeline, cost models + potential dependencies- Laurence Livermore
   ● Single milestone, single deliverable
   ● Lots of inter linkages with other work packages
   ● ICEDIG, Synth and Dissco tasks
   ● Costing and understanding services - understanding who the users are
   ● WP3 3.3 - formal staff secondment procedure
   ● Not just technical expertise, but neutral third party for audit/quality assurance and control

4) Proposed subtasks:
   ● Four main points (see slides):
   ● Digitisation costs - three categories under costs; mobilization (physicalities), data processing + data access & exploitation
   ● Possibility of data hosting - key component
   ● Expanding scope of ICEDIG report, integrating with overall cost model
   ● IT services necessary for operation for DiSSCo community (internal and external comms).
   ● Proposal to store costs in a sustainable way - web-based application- by technical sub-contractor
   ● Neutral third party to check user needs of DiSSCo - by technical sub-contractor

5) Agree Next Steps
   ● Delimit scope of task 4.1 & 4.2
   ● Read and discuss recommendations from ICEDIG reports
   ● Review links and dependencies with other WP
      ○ Scope of person months; operational cost operation
   ● Discuss Cost Model platform with DiSSCo Techgroup
      ○ Discuss with DiSSCo group; long-term maintenance

Discussion:

   ● LL: Previous suggestions by Pat for Financial elements.
   ● PM: Actually more linked to 4.4. - establishing pre-commercial procurement procedures to deal with services. Contact with a consultant who did similar work.
   ● PM: Outcomes of 4.4, and 8.3
   ● LL: Useful to have more detail on procurement, financial expertise - EOSC (European Open Science Cloud) cost modules.
   ● PM: **Cost depends on whether it is for private parties or provided to private companies.**
   ● LL: How to capture cost information for this model, "How to price a collection for digitisation". Tendering process for digitisation.
   ● LL: Picturae - scale model used for European collections compared to ? model for Smithsonian. Worth thinking about how the Smithsonian model is used?
   ● AC: Referring to work package 5 - Difficulties to liaise with the private sector - not open to co-create and develop. Currently redesigning the scope for this

task due to these challenges (from w/b 25th Jan), with a deadline for submission, 31st March. Necessary to create a Plan B if number of submissions not met.

- LL: What should be included in 4.1 & 4.2 (details of digitisation in latter, granular approach, building on parameters from 4.1). What else should be considered to separate the tasks.
- MG: 4.1 is the singular steps 'the bricks'; 4.2 'assembles the elements from 4.1
- LL: Task 3 will consider precuration costs. Group tasks by Nelson et al's five task clusters. Physicalities of collection must be noted before digitisation project commences.

Thoughts on proposed subtasks - Laurence Livermore:

- EA: Need better structure, more points for deliverables. How to articulate milestones.
- LL: Most concerned about platform for storing costs. Worth revisiting technical subtractor, especially in reference to DiSCCo tech team.
- MG: Need to identify exact needs. One month to identify, for them to start in the summer. - better to take the time and find someone experienced than have to go through the process again.
- PM: Contractor who did work made interesting adaptations and suggestions, used in application towards 4.1.
- LL: Worth 'sounding them out'. More specialist expertise needed.
- EA: Have the money required to do this. How to use services up to you. Amendment only if increase in subcontractor funds/transfer of money between work packages.
- LL: Patricia, in terms of your institute's involvement - how do you feel about work package four?
- PM - *mic distortion*
- LL: As task leader, any guidance on NDAs?
- EA: Will check and send the communication. There are prepared guidelines.
- LL: To Quentin - Thinking about structuring, enriching, enhancing data and developing that into a service - can this be integrated in the work package, especially concerning pre-digitisation costs.
- QE & S: Combine both costs together.
- LL: What is the pre-digitiation to publication timeline compared to Picturae.
- S: Have two projects to compare to internally.
- EH: - DWG (digi working group) focusses on minimal info for specimens, minimal info for images - specifications have implications for costs.
- AC: Tasks not specifically intended to reference to service costs. What are the components of the costs that need to be considered for building services as a unit cost. Addressed in WP4 or a different task?
- MG: 4.1 = the bricks, 4.2 = the assembly!
- AS: **4.2 should integrate the elements and define the costs of services**
- LL: Real service cost vs. presumed cost. Does cost change depending on provision to different user groups.

- ○ EF: Project has a business model umbrella, indicator costs should be associated with specific activities and users. Possible combination of costs and incomes later.
- ○ LL: Included capture of full economic costs, but how would profit be determined via these services. Example of Synth. Would full economic costs be useful?
- ○ AC: Need the cost basis before thinking of charges, margins or benefits. This falls under the business model which DiSSCo hasn't yet fully acknowledged and needs to take into consideration.
- ○ MG: Connection between costbook, service and business model related to 4.3 (funding by government). Digitisation services may have partly already been paid, so charges depend if gov partly pays for this (more info on this?)
- ○ EF: **Costs and benefits; not strictly financial, but also societal (culturally etc.) Related to 1.4; these points must be linked.**
- ○ LL: The societal impact relates to the end process of digitisation.

Key discussion - Milestone report:

LL: Overdue, summary of subtasks and milestones, discussion. 3-4 page draft, for end of the month.

EA: Needed for all task leaders. Quality of outcomes should be priority. Need to be aware of dependencies, your task should not jeopardise work of others; be conscious of others.

LL: Not so confident in dependencies of tasks other than 4.2. Clarify scope of relationship by end of Feb; planning organisation of subcontractor by end of Feb.

PA: Document produced with Salome of requirements and needs. Good to have LL contribute to this by the end of Feb.

Key discussion - Meeting with DiSSCo technical team:

LL: Early Feb meeting to discuss platform. Good to have someone from Paris to contribute to this.

MG: Depends on work package 6

CW: Current services in AWS (Amazon Web Services). Scope to support some WP4 needs in WP6.

LL: How would data be captured and managed in a more sustainably, standardised format. Meeting to be set with CW and Sharif(?) by deadline (end of Feb).

Other:

Clarification from EA on dependencies:

EA - OK to clarify dependencies by end of Feb, including sub tasks. No need to change deliverable dates.

End of discussion, 13:15.

Notes from chat

From Elspeth Haston to Everyone: 11:59 AM
Nice cats!

From Patricia Mergen to Everyone: 11:59 AM

one more ;-)

From Michel Guiraud to Everyone: 12:00 PM

Name is Canelle!

From Patricia Mergen to Everyone: 12:00 PM

mine Shessomaru (Killing perfection)

From Elspeth Haston to Everyone: 12:05 PM

Had to google translate to get "Cinnamon" :)

The two ICEDIG deliverables are not yet on the ICEDIG website.

From Elsa Fontainha (ULisboa) to Everyone: 12:06 PM

WP 1 received by email the d 2.1. and 2.2

From Eva Alonso to Everyone: 12:07 PM

You can find them in teamwork

From Elspeth Haston to Everyone: 12:08 PM

@Eva Is that the ICEDIG teamwork?

From Eva Alonso to Everyone: 12:08 PM

Yes

From Elspeth Haston to Everyone: 12:08 PM

Ah. Not in ICEDIG so don't have access

From Eva Alonso to Everyone: 12:11 PM

You need to have access to the project but once there, you have all in Files

I see the D2.2 there. However, I do not manage ICEDIG only have access, so you will need to talk with Luomus

From Mathias Dillen to Everyone: 12:12 PM

d2.2 is on the icedig website; d4.5 and 8.2 aren't.

From Elspeth Haston to Everyone: 12:13 PM

I was looking to see if D4.5 and D8.2 were available to people.

From Mathias Dillen to Everyone: 12:13 PM

https://riojournal.com/article/56211/download/pdf/

https://riojournal.com/article/58915/download/pdf/

From Elspeth Haston to Everyone: 12:14 PM

Would be good to get these on the ICEDIG website and on the DiSSCo site here

https://www.dissco.eu/what-is-dissco/knowledge-base/

From Mathias Dillen to Everyone: 12:15 PM

and on zenodo (at least in the icedig community there)

From Elsa Fontainha (ULisboa) to Everyone: 12:15 PM

@Elspeth thanks.

From Elspeth Haston to Everyone: 12:16 PM

@Mathias - thanks for providing the links here

From Me to Everyone: 12:20 PM

Minutes/Notes:

https://docs.google.com/document/d/1Hb5M5JwtXoiOIAvyz6QapalNAlH7-y_HPdJgq1tlRcY

https://docs.google.com/document/d/1_KSfJRSX7zmDoo3qVbWI2S0fUrfaa8XuwmbBXTvu9L8

Task   4.2   proposal   notes:

https://docs.google.com/document/d/1_KSfJRSX7zmDoo3qVbWI2S0fUrfaa8XuwmbBXTvu9L8

From Elsa Fontainha (ULisboa) to Me: (Direct Message) 12:24 PM

The 'real'costs are combined with the actual 'cost of services'. (similar with sme services in public hospitals "you pay x but the real cost is >x)

My question is about the management of financial results of different uses (and users). The cost indicators will be associated with charged costs (is my question). All costs must be included (human resources, capital costs, current costs, etc.). Is not easy this allocation for each specific service .

From Ana Casino to Everyone: 12:32 PM

This is the link to SYNTHESYS+ NA5.3 Call for proposals to the project's partners

https://docs.google.com/document/d/1MVDaN-vAEkXm-30G4ov0QjFPq3bvCWD3/edit#heading=h.gjdgxs

From Francois DUSOULIER (MNHN) to Everyone: 12:41 PM

@Pat You have a problem with your microphone

From Lizzy to Everyone: 12:41 PM

@Pat, sorry I can't quite hear what you're saying!

From Patricia Mergen to Everyone: 12:44 PM

I have to log out and log in again,when it does that

the experience with the call for tender with Picturae ….

From Elsa Fontainha (ULisboa, Lisbon School of Economis and Business) to Me: (Direct Message) 12:50 PM

@Ana is partially answering my question.

From Elspeth Haston to Everyone: 12:55 PM

I'm afraid I have to leave. Thanks very much for the session, Laurence. Will be in touch about links with CEGAF DWG and MOBILISE WG2.

From Me to Everyone: 12:55 PM

Thanks Elspeth

Session Task 5.1 "DiSSCo Knowledgebase"

Session outcomes:

- Working session: around 40 participants
- introduction and demonstration of first draft version of DiSSCo Knowledgebase using the open source software package DSpace 6
- positive feedback from participants on functionality and progress made so far
- participants appreciated being involved in decision making
- discussion resulted in prioritization of how to proceed
- next steps should focus on User Interface (usability, data structure, findability, submission workflow) instead of technical refinements (like API) which would follow in a later step

Session Task 5.1 "DiSSCo Knowledgebase"

Agreements and next steps:

- improve layout of user interface (check existing style recommendations for corporate design)
- implement session feedback regarding user interface; for example ideas for:
  - how to guide users on the homepage,
  - how to structure submitted data
- implementation of authentication mechanism (OpenID, Single sign-on SSO)
- internal launch of DiSSCo KB: planned for February 2021
- set up GitHub for collection of issues
- productive launch after testing and improving by DiSSCo partners
- automatic DOI assignment: second half of 2021

1. Session: WP5, Task 5.1 DiSSCo Knowledgebase

January 19, 2021 | **Convener:** Sabine v. Mering / Julia Pim Reis / Falko Glöckler / Mareike Petersen

Around 40 participants attended the working session of Task 5.1 of WP5. Participants included members of all institutions actively involved in the task but also members from other work packages and the Coordination and Support Office (CSO).

List of participating institutions:
Botanic Garden and Botanical Museum Berlin, Freie Universität Berlin, Germany
Cardiff University, UK
Meise Botanic Garden, Belgium
Museum für Naturkunde Berlin, Germany
Musée national d'histoire naturelle, Luxembourg
Muséum national d'histoire naturelle, Paris, France
National Museum Prague, Czech Republic
Natural History Museum London, UK
Natural History Museum, University of Florence, Italy
Naturalis Biodiversity Center Leiden, Netherlands
Royal Botanic Garden Edinburgh, UK
Royal Belgian Institute of Natural Sciences (RBINS), Brussels, Belgium
Senckenberg (SGN), Frankfurt am Main, Germany
University of Lisbon (ULisboa), Portugal

2. Main discussion highlights and decisions

The aim of the session was to introduce the first session of the DiSSCo Knowledgebase and to collect feedback from the participants with a focus on the user interface. Further requirements and priorities for the development of the DiSSCo Knowledgebase were to be discussed.

Session outcomes:
- introduction and demonstration of first draft version of DiSSCo Knowledgebase using the open source software package DSpace 6,
- positive feedback from participants on functionality and progress made so far,
- participants appreciated being involved in decision making,
- organizational structure and collections of related information/documents should be improved,
- structure of Knowledgebase needs to allow growth (additional projects or DiSSCo core topics),
- discussion of several topics which need to be discussed by the DiSSCo Technical Team, beyond DPP (e.g. identifiers / usage of DiSSCo DOIs),
- discussion resulted in prioritization of how to proceed,
- next steps should focus on User Interface (Single sign-on SSO, usability, data structure, findability, submission workflow) instead of technical refinements (like API) which would follow in a later step.

3. Next steps

Agreements and next steps:
- improve layout of a richer and more functional user interface
  (check existing style recommendations for corporate design)
- increase amount of content (e.g. ICEDIG Deliverables, Milestones (check if still relevant and check with authors on public or project-internal availability)
- implement session feedback regarding user interface; for example ideas for:
  - how to guide users on the starting page
    (e.g. "Are you new here?" "Are you interested in DOIs?")
  - how to structure submitted data
  - tagging of content
- implementation of authentication mechanism (OpenID, Single sign-on SSO)
- internal launch of DiSSCo KB: planned for February 2021
- set up (private/internal) GitHub Repo for the collection of issues and suggestions
- productive launch after testing and improving by DiSSCo partners
- automatic DOI assignment: second half of 2021
- The Knowledgebase was mentioned as a possible home for various project outcomes during the AHM. There will be a close collaboration for e.g. Policy Framework (Task 7.3) and Criteria to identify a prioritisation model for digitisation (Task 1.3).

4. Conclusions

This working session during the first All Hands Meeting proved very useful for the ongoing process of developing the DiSSCo Knowledgebase. Considerable progress was made and further developments discussed. The feedback collected from project partners will help to further develop the first draft version and to meet the requirements specified by the project partners. The next steps and priorities defined for the work within Task 5.1 will move us closer towards the final implementation of the DiSSCo Knowledgebase.

5. Reference documentation

Petersen, M., Pim Reis, J., von Mering, S. & Glöckler, F. 2020: The DiSSCo Knowledgebase. Blog post published on the DiSSCoTech blog on December 18, 2020. Online at:
https://dissco.tech/2020/12/18/the-dissco-knowledgebase/

https://www.dissco.eu/what-is-dissco/knowledge-base/

Outcomes and actions from TASK 5.4 meeting:

"Modernising technical infrastructure for science data mobilisation and publication"

Friday, January 22nd 2021

Lead: Matt Woodburn (NHM London)

January 22nd 2021

TASK OVERVIEW AND MEETING SUMMARY

Deliverable 5.5 - Construction plans for the improvement of technical infrastructure in the areas of geo-collection data and taxonomic services

Task starts on 1st April 2021, deliverable due July 1st 2022

Meeting topics

Presentations on:

> DiSSCo services and architecture (Wouter Addink)

> GeoCASe 2.0 progress and roadmap (Falko Glöckler)

> Catalogue of Life progress and roadmap (Olaf Banki)

Follow-up discussion about the task scope and focus, and the bigger picture of interactions between DiSSCo and services

AHM1

Task start

Deliverable 5.5

April

2021   **2022**   **2023**
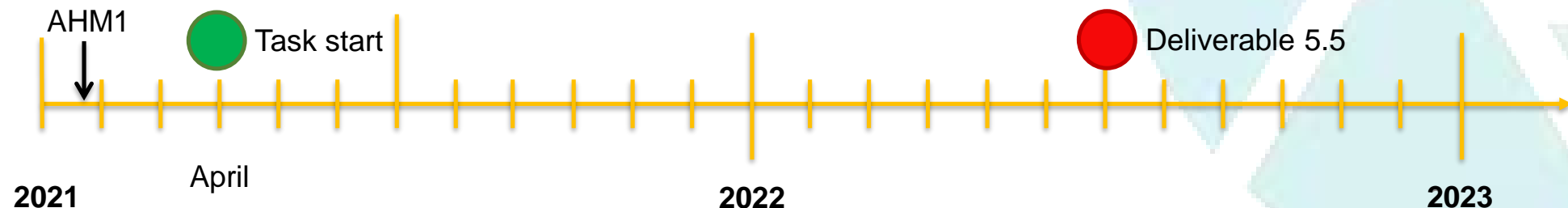
OUTCOMES AND NEXT STEPS

- DiSSCo, GeoCASe and Catalogue of Life are all moving rapidly ahead in modernising their technical infrastructure. However, the latter two will need construction plans for integrating with DiSSCo, and vice versa.

- DiSSCo will incorporate at least 5 services (ELViS, ECAS, SDR, Knowledgebase and collections monitoring dashboard), which will add complexity to the interactions.

- An initial focus of the task should be to start mapping these interactions, towards a technical blueprint for DiSSCo data integrations.

- Suggested approaches include scoping through small pilot activities, and taking an event-focused approach to mapping interactions.

- **Next steps -** follow-up meeting of task partners to consider:

  - (internal) milestones and subtasks
  - overlaps with other DPP and SYNTH+ tasks, DiSSCo Tech etc

DiSSCo Prepare All Hands Meeting - Task 5.4

Update on GeoCASe and CoL+ roadmaps, and discussion of subtasks and partner roles

Thursday 21st January 2021, 10.45-12.15 CET / 09.45 - 11.15 GMT

Meeting agenda and notes:

https://docs.google.com/document/d/1JwJ90ru3eDRM0NpMsYQqMv_XsmBiboNHx0YkGoDGaZs

Main discussion highlights and decisions

Presentation - DiSSCo update (Wouter Addink)

- DiSSCo currently includes 5 services either planned or in development: ELViS (European Loans and Visits System), ECAS (European Curation and Annotation System), collections monitoring dashboards (including the pilot Collections Digitisation Dashboard), the SDR (Specimen Data Refinery) and the DiSSCo Knowledgebase.

- These services will (to different degrees and in different ways) connect into the core DiSSCo architecture, and other systems e.g. CETAF

- DiSSCo is also part of BICIKL (Biodiversity Community Integrated Knowledge Library), a 3 year project starting in 2022 to building a community of key research infrastructures in the domain of biodiversity

- A DiSSCo digital specimen architecture pilot, supporting the openDS specification, will run from 2021-2022, starting in February 2021. Two other pilots (CMS integration and Specimen Data Refinery) are also planned.

- Challenges for DiSSCo with respect to integration with external services include:

  - CoL getting data from DiSSCo

  - Resolvable persistent identifiers for CoL names and taxon concepts

  - Showing different opinions and names from different name providers (e.g. NCBI taxonomy)

  - Resolvable PIDs for GeoCASe

  - BioCASe and IPT

- ○ Governance and funding model for GeoCASe

- ○ Combined progress indicators for data shared from GBIF and GeoCASe

- ○ DiSSCo user stories that need implementation in GeoCASe or CoL

- ○ Digital Specimen data in GeoCASe beyond ABCD/EFG

Presentation - GeoCASe 2.0: The Earth Science Collections Portal (Falko Glöckler)

- ● Background to GeoCASe:

  - ○ GeoCASe v1 was built by MfN in 2007 within the SYNTHESYS project, to fill a gap for aggregating geological object data. The service hosts data for both geology fossil specimens.

  - ○ The EFG extension to the ABCD standard was developed and implemented for the first version of GeoCASe.

  - ○ There is overlap between GeoCASe and GBIF in the hosting of palaeontology object data.

  - ○ The CETAF Earth Science Group (ESG) is highly engaged in supporting GeoCASe.

  - ○ Some more development of GeoCASe was carried out at MfN in 2016, and in 2020 a new prototype was developed by the Geology Department at Tallinn University of Technology (TalTech)

  - ○ GeoCASe 2.0 is under active development at MfN, with a soft launch planned at the end of January 2021. A preview can be seen at https://new.geocase.eu

- ● GeoCASe 2.0 has a modernised architecture, new interfaces and a large number of new features, including a new REST API.

- ● The GeoCASe roadmap:

  - ○ Phase 1 (2020) is to replace the old portal, nearly complete.

  - ○ Phase 2 (2021) focused on improving data quality, including data harmonisation and enrichment, machine readability and data linking. It will also focus on engaging data providers to source more data.

  - ○ Phase 3 (2022) is a construction plan for integrating into the DiSSCo RI.

Presentation - The new Catalogue of Life: A key taxonomic service for DiSSCo (Olaf Bánki)

- Through the CoL+ project, a new CoL infrastructure is in place, powered by GBIF. It came online in December 2020.

- Features include:

  - The CoL ChecklistBank (https://data.catalogueoflife.org), an open repository to publish checklists and nomenclatural datasets. The new infrastructure has enabled it to scale up to expected usage, moving away from manual operations and be more adaptable for future use cases.

  - The new CoL portal (https://catalogueoflife.org) provides access to checklists but is also integrated with other information about the organisation, partners etc. It has a variety of new features, including taxon pages with better linkage and source references, a browsable global species checklist tree, and better search. It's built using React so that modules can also be reused on other sites. Old web services, access to previous CoL checklist editions in Docker containers and DwC-A/MySQL downloads of annual editions are also still supported.

  - The new CoL API (https://api.catalogueoflife.org) provides the backend for the portal. Development is continuing in coming months, including work in progress on stable taxon name identifiers, and a potential new data standard for CoL data packages (into which community input would be useful). It is however still supporting DwC-A as well as ACEF. The legacy API is still available but will be deprecated and is not recommended.

- Data comes from data publishers (e.g. Taxonworks, WoRMS), often through automated processes. Issues and metrics are made available to publishers.

- Data is transferred into the checklist with some interpretation and merging. The tools allows editing of the draft before publication.

- Next steps for 2021:

  - Consolidating the infrastructure e.g. connections with data publishers

  - Extending CoL to replace the GBIF backbone taxonomy

  - Adding accreditation services e.g. DOIs, Open Data licenses, better statistics and feedback mechanisms

  - Building a global vision as a foundation infrastructure for species names, including ongoing discussions with iDigBio

- Strong interactions between CoL and DiSSCo already, via CoL steering committee, SYNTHESYS+ (SDR), DiSSCo Prepare WP5, and BICIKL.

- CoL needs to do more to support taxonomic services for fossils. A working group has been set up, and invitations will be made to key partners e.g. PBDB.

- The discussion on how CoL interacts with DiSSCo services needs to begin. Interactions are likely to be at a number of different scales, for example at institutional level, with the SDR, and with openDS.

Open discussion

- Should DiSSCo take data from the data aggregators, or from data publishers (institutions) directly? What should the relationship between the three be?

  - GBIF and GeoCASe already have provided pipelines for data collection from publishers. DiSSCo could use that, but aggregators could also provide additional data to DiSSCo as they do data cleaning and interpretation, which should be provided to DiSSCo as annotations.

  - For the verification of enriched data, aggregators like GBIF and GeoCASe could take a role in orchestrating the feedback loop to data publishers.

  - Mechanisms will be needed to make sure that everything is in sync.

  - There is a triangle between collections, aggregators and DiSSCo - could DiSSCo perhaps have a different role over the top? What is the relationship between DiSSCo and aggregators?

- How easy would it be to create a purely technical blueprint for interactions between these platforms, without taking into account other influencing factors e.g. cultural, political?

  - These can't be completely divorced. We'll therefore need a way to capture those non-technical requirements from appropriate groups and sources.

  - CoL interactions with DiSSCo will be different to those with other aggregators, and we need to look at the relationship at the right levels to define this, and understand the landscape.

  - The content strategy also needs to be clear, e.g. knowing where the data is cleaned and enhanced.

  - DiSSCo architecture has a number of components which will likely link with CoL and aggregators in different ways and at different levels.

- What approaches can we use for these challenges?

  - Interactions can be scoped through small pilots to start understanding the issues, and produce a concrete roadmap.

  - A lot of coordination is required between projects, due to a lot of overlap.

- ○ We can start by discussing at a meta-level, for example thinking of it like an event bus between services - what happens when x happens in y and what event is pushed across to other places. Then, we can think less about what needs to be done with the event and more about what events we need.
- What are some main challenges with the meta level thinking?
  - ○ How the aggregators and DiSSCo architecture will interact.
  - ○ How to get additional information from aggregators about the digital specimens, on top of the existing data.
  - ○ How to match between the records in the aggregators and DiSSCo, as DiSSCo is likely to need to use a combination of the aggregators data and DiSSCo's data, and for example both GBIF and GeoCASe have fossil data.

Next steps

Next steps are to arrange a follow-up meeting of task partners, possibly prior to the start of the task on April, to:

- follow up on conclusions and suggestions from the AHM session
- draft internal milestones and subtasks for T5.4
- discuss overlaps with other DPP and SYNTH+ tasks, DiSSCo Technical Team work, and initiatives outside of DiSSCo projects

Conclusions

- DiSSCo, GeoCASe and Catalogue of Life are all moving rapidly ahead in modernising their technical infrastructure. However, the latter two will need construction plans for integrating with DiSSCo, and vice versa.

- We're collectively yet to understand what those interactions should be, and we need a better understanding of that before we can move onto the question of how they could be technically implemented.

- DiSSCo will incorporate at least 5 services (ELViS, ECAS, SDR, Knowledgebase and collections monitoring dashboard), which will add complexity to the interactions.

- Interactions are also going to vary for different types of external services, for example an aggregator like GeoCASe is likely to interact with them in different ways to a service like the Catalogue of Life.

- There is a triangle of DiSSCo, external services and institutional CMSs, and we need to figure out through which pathways and in which direction data needs to flow between them. There is an obvious link into the work planned in Task 6.1 on CMS systems interoperability.

- An initial focus of the task should be to start mapping these interactions between DiSSCo components and external services, with Catalogue of Life and GeoCASe as primary use cases, and incorporating CMS interoperability in conjunction with T6.1.

- Potential approaches include:

  - scoping through small pilot activities, which may feed into planned pilot and demonstrator activities already being planned within DiSSCo

  - taking an event-focused approach to mapping interactions, so that we can think about interactions at a meta level, without simultaneously worrying about the routes and mechanisms by which they would occur.

Reference documentation

- GeoCASe portal: http://www.geocase.eu/

- GeoCASe Github repository: https://github.com/MfN-Berlin/geocase-infrastructure

- Catalogue of Life: https://www.catalogueoflife.org/, https://data.catalogueoflife.org, https://api.catalogueoflife.org

- Catalogue of Life repository: https://github.com/CatalogueOfLife/general

Session recordings

- Whole session: https://drive.google.com/file/d/1mPt5BRHLSdB4lWYcHLLUK8w5zVFxKvvj

- Presentations:

  - DiSSCo: https://drive.google.com/file/d/15DS_f0cihWzPY2mdYdAACsjpLA8yxwCu

  - GeoCASe: https://drive.google.com/file/d/1_03w9_KkpjT1-IriWqbXeEFfZiYY_DEz

  - Catalogue of Life: https://drive.google.com/file/d/1nZs9d4voDDws1gWh8NYVWzjaQY7L5PWa

openDS I+II keynotes & discussion

- **openDS**
  - standard specifying and detailing **DiSSCo's data model DIgital Specimen**, a digital twin of a physical specimen in a collection
  - **key cross-cutting topic**, essential for all tasks in WP6 "Technical Architecture & Services provision" (is based also T5.2)
  - addressed by **keynotes and open discussion** on progress and gaps
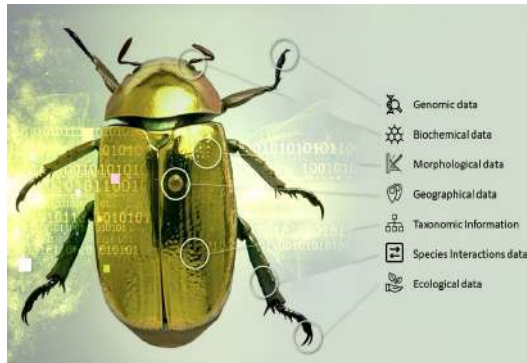
Outcomes T6.1 CMS interoperability

- Progress: connecting physical and Digital Specimen with regards to standardization (**minimal information model MIDS**) and reference platform (**open CMS DINA**)
- Challenges: High demand for "CMS as a Service" + Synchronization between local CMS & central DS architecture critical
- Workplan: Evaluate DINA's microservice architecture as **CMSaaS**. + Assign **PID(s) for specimen at earliest point possible** (links to MIDS)
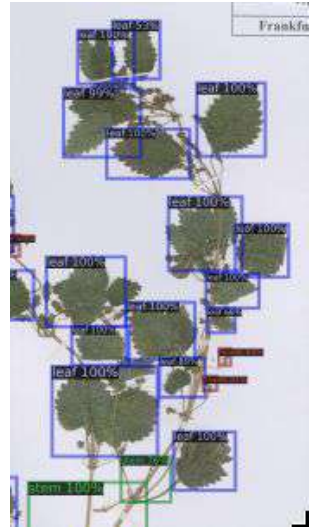
Outcomes T6.2 Digital Specimen architecture

- Progress: **Adaptation study of FAIR Digital Object architecture** for DiSSCo based on RDA recommendations http://doi.org/10.5334/dsj-2020-050
- **Convergence of Digital** (DiSSCo)**/Extended** (BCoN) **Specimen** approach
- Challenge: Again: Synchronization CMS, DS architecture and linked external resources (e.g. sequences in ENA)
- Workplan: **Elaborate Digital Specimen repository** (Cordra-based) + Link Builder
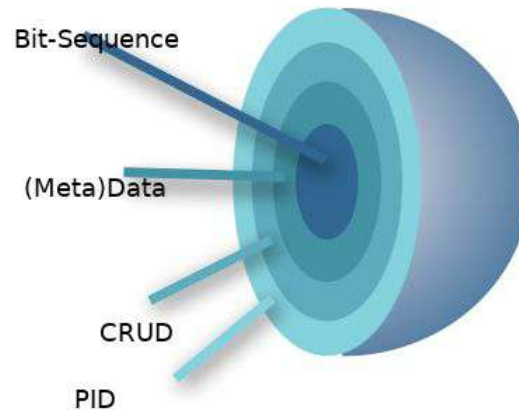
Outcomes T6.4 Embedding in the technical landscape

- Progress: Intensive **interaction with Elixir** (Biohackathon project), **EOSC** (SEMAF - semantic mapping study), **iDigBio/BCoN** (ES/DS convergence, NSF application for Digital Object store for ES prototype)
- Challenges: **DwC/ABCD mappings** and "adjacency" to OBO ontologies **crucially** for openDS
- Workplan: GBIF and EMBL provide a new dataset enabling consolidation of the **voucher/seq linking**
  + Contribute to **semantically richer DwC dictionary** mapped to OBO ontologies
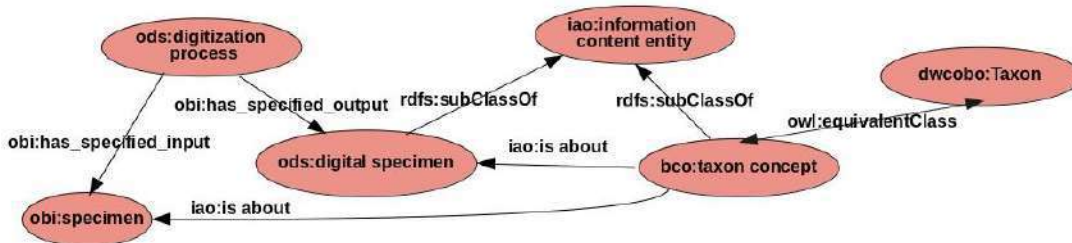
DiSSCo Prepare AHM1

openDS sessions

This googledoc:
https://docs.google.com/document/d/1XD3EyqGbISIZmfy9yJJRyUTQx7gwBupg0Uio9Rusr6c/edit?usp=sharing

Agenda (again)
https://docs.google.com/presentation/d/10wpwi8tNpIWzu8_Re20uITBqare9JA3CS00JfzKFxBg/edit?usp=sharing

Session Type: Brief presentations and discussions

Date: Wednesday January 20., 9:00-12:00 CET

Organizers: Claus Weiland, Senckenberg
        Alex Hardisty, Cardiff
        Wouter Addink, Naturalis
        Sharif Islam, Naturalis

Notes: Hanieh Saeedi, Senckenberg
        Thomas Winter, Senckenberg
        Jonas Grieb, Senckenberg

Aim of the session:
We will use the double session to look at the DPP WP6 work more widely to examine and reflect our progress concerning deliverables, milestones as well as general direction and outcome of the WP "technical architecture & service provision". The format of the workshop is an open discussion guided by keynote presentations and should lead us to detect gaps and make suitable adjustments to our work program.

Recommended Material:
Glöckler et al. (2020). DINA - Development of open source and open services for natural history collections & research. https://doi.org/10.3897/biss.4.59070
https://github.com/tdwg/mids
https://github.com/dissco/opends
https://github.com/elixir-europe/BioHackathon-projects-2020/tree/master/projects/33
Recorded video on ES/DS comparison and technical approach:
https://www.dropbox.com/s/ajvxqf6os2d4gzz/Archi-v0.4-16Dec2020-export.mp4?dl=0
Islam, S. et al (2020) Incorporating RDA Outputs in the Design of a European Research Infrastructure for Natural Science Collections. https://doi.org/10.5334/dsj-2020-050
Raes, N. et al (2020) White paper on the alignment and interoperability between the Distributed System of Scientific Collections (DiSSCo) and EU infrastructures - The case of the European Environment Agency (EEA). https://doi.org/10.3897/rio.6.e62361
Lannom, L. et al (2020) FAIR Data and Services in Biodiversity Science and Geoscience. https://doi.org/10.1162/dint_a_00034

Agenda:
9.00-10:30 Session Tasks 6.1 + 6.2
Intro - Claus Weiland(mod)/Hanieh Saeedi(notes)
Introduction to the ideas of DINA and it relation to DiSSCo - Falko Glöckler/MfN
Introduction to MIDS - Elspeth Haston/RBGE
[Discussion gaps/adjustments T6.1]
Comparing Digital Specimen and Extended Specimen Concepts - Alex Hardisty/Cardiff (recorded video, 12 minutes for pre-viewing above)
Fair Digital Object and RDA output incorporation in DiSSCo's design - Sharif Islam/Naturalis
[Discussion T6.2]
[15 min break]
10:45-12.15 Session Tasks 6.3 + 6.4
Continuation - Hanieh Saeedi(mod)/Claus Weiland (notes)
[Continuation discussion T6.2 and discussion T6.3, which has yet to start]
[Connecting molecular sequences to their voucher specimens](#) - Mathias Dillen/MeiseBG
[Integration of biodiversity data into EOSC through a flexible semantic mapping framework](#) - Claus Weiland, Senckenberg
Digital convergence and global specification for data integration - Alex Hardisty/Cardiff
[Discussion T6.4 and WP6]

Participants:

Niels Raes - Naturalis
Sharif Islam - Naturalis
Wouter Addink - Naturalis
Anton Güntsch - Freie Universität Berlin, BGBM
Sarah Rossi de Gasperis (NHMuseum, University of Florence)
Pieter Huybrechts (Meise Botanic Garden)
Jiri Frank (National museum, Prague)
Josh Humphries (Natural History Museum, London)
Matt Woodburn (Natural History Museum, London)
Maarten Trekels (Meise Botanic Garden)
Jose Alonso (Naturalis)
Judite Alves (Museu Nacional de História Natural e da Ciência, Ulisboa)
Sabine von Mering (Museum für Naturkunde Berlin)
Philippe Loret (Muséum National d'Histoire Naturelle, Paris)
Pierre-Yves Gagnier (Muséum national d'histoire naturelle à Paris)
Laura Tilley- CETAF
Mikko Heikkinen - Luomus
Ville-Matti Riihikoski - Luomus
Esko Piirainen - Luomus
Lorenzo Cecchi (NH Museum, University of Florence)
Gianna Innocenti (NH Museum, Florence University, Italy)
Mathias Dillen (MeiseBG)
David Fichtmueller (Botanic Garden and Botanical Museum, Berlin)
Patrick Semal (RBINS)
Eva Alonso (Naturalis)
Patricia Mergen (MeiseBG)

Ana Casino (CETAF)
Julia Pim Reis (Mfn, Berlin)
Maria João Fonseca (Natural History and Science Museum, U.Porto)
Tina Loo (Naturalis)
Quentin Groom (Meise Botanic Garden, Belgium)
Elspeth Haston (Royal Botanic Garden Edinburgh)
Inês Pinto (Museu Nacional de História Natural e da Ciência, Ulisboa)
Sofie Meeus (Meise Botanic Garden, Belgium)
Wesley Tack (Meise Botanic Garden)

Notes:

- Falko: Presentation of DINA consortium
- Why DINA: Diversity of stakeholders (curators, researchers, software developers) lead to different requirements, can't be covered in a monolithic system.
- Problem alignment of different tools
- Important challenge: aligning the existing tools
- Challenges regarding the data management (storages, specimens, research data, ...external research data and authority data, connection to other collections)
- https://github.com/DINA-web/guidelines
- DINA approach: incorporating the component and create a interoperable component through a web through three ways, 1-development, 2- refactoring, 3- wrapping (incl software tools)
- DINA approach: different components are loosely coupled with web APIs -> microservice architecture, take from DINA what you need as plugin
- The challenge of redundant development: the key is to have a join development, this component can be implement in other systems
- Distributed development of components, key: agreement on APIs
- Challenge of harmonization: what should be the functionality of a connector to the system, we need to cope with 1- physical object (digitisation, additional data, changes in the collection, digital request such as loans and the limitation for the digitisation), and 2-
- Digital specimens are not linked to the physical object and can be reused in research
- Physical events evoked by digital events (Digital information might need physical verification)
- How do we cope with linkages between the data and DiSSCo?

Questions:

- Q: How can people contribute to DINA
- A: everything is publicly online, the drafts and codes, can be found on github, you can be an associate member, for becoming a core member require dedicated resources, Dina require support by dedicating resources from groups who are working on collection management and in this case you will become a core member, DINA invite the institution to align with DINA and contribute in coding
  https://www.dina-project.net/wiki/Welcome_to_DINA
  https://github.com/DINA-Web
  Q: how to cooperate with DINA
  A: you need to check the security online, you need to make your source codes publicly available and should be a community driven
  Q: are there guarantee from DINA for quality of the code

A: DINA core members agreed to sign a maintenance contract, it depends on the component

Note from Patrick Semal: The risk is the same with a commercial company ...

Note from Josh Humphries: a point on the security of open source projects - the benefit of security issues being spotted in open source projects is only relevant if there are lots of eyes looking at the code.

Q: long-term problem is contribution the developers into data infrastructure, what CMS will you recommend for users, we need more integrated system, is it possible to have a central DINA

A: Technically yes, but depends on the components

Q: Is it more like that in a CMS, can we get the resources that take multiple tenants?

A: Yes, that would be a good option.

Q: At what extend the help desks are at DINA, do what extend DINA exchange data with GBIF

A: there is no exchange with GBIF as ther are working in different component, DINA would use IPT to publish data to GBIF, there is no help desk at DINA yes, the current approach is getting part of the story, if the component will extend the story might change

Q: Who hold the responsibility of keeping data, do you want GBIF to do it

Q: Why Cordra? When and how was that decision made?

- Elspeth Haston/RBGE: Introduction to MIDS -
- How to connect the physical and Digital Specimen
- There are three core levels in standardisation the linkage, 1- basic (basic record of specimens), 2- regular (partial data records, where key information is stored), 3-extended (other data present the specimen and third-party resources)
- DS aims to make the data publicly available and accessible, expect the minimum level of digitisation
- Basic level: two main groups, 1-TDWG and 2- CETAF digitisation working group
- Some of the element might not be included in the final standard, but the recommendation of the data will be included
- MIDS: Focus is data presence, not quality (which is of course an issue)
- The issue with missing data (e.g. sensitivity issues, and other reasons, the aim is to push people to fill the missing data, and consistent use of terminology for missing data)
- Regular level: this level include the geographical information, andy type data
- Extended level: This include the determination, and links to other information
- If you want to contribute to the TDWG Task Group you can sign up to its mailing list here: http://lists.tdwg.org/mailman/listinfo/tdwg-mids
- Note from Pieter Huberechts: +1 Elspeth, I stress that the capture of missing data is very valuable
- TDWG collection description Interest Group and OpenDs,
- Note from Addink: There is a DwC field for institution referent, which is dwc:institutionID. Issue with that field though is that it includes a recommendation to use an identifier from a collections registry, and cannot capture the type of identifier (GRID, ROR etc)
- DwC needs to implement the needs for DS
- Note from Quentin: the process of implementation for DS is rather slow
- Note from Alex Hardisty: it is important for open DS to build on DwC, if DiSSCo can take the lead on this? The necessity of training people to get the knowledge to do this

- Challenge with the images: complexity of the images which slow down the process of standard digitisation, the approach is developing the tas of extension of images and identification
- DINA -> CMSaaS due to microservice architecture
- IPT connected to the core DINA
- Niels: Helpdesk services
- Alex Hardisty/Cardiff: Comparing Digital Specimen and Extended Specimen Concepts
- WP6 technical background presentations: http://bit.ly/DiSSCoWP6Share
- Explanatory video, 12 minutes: https://bit.ly/esdsframework
- Positioning openDS in the landscape: https://github.com/DiSSCo/openDS/blob/master/positioning-opends.md
- To follow the opends work: https://github.com/DiSSCo/openDS
- Alex: Talk on ICEDIG plan leading to implementation and construction plan
- Conceptual scope: infrastructure for Europe
- The important thing was how to mage the data in DiSSCo, recommendations on tools, but there was no plan, thus task 6.2 took over
- The last 12 month, was the development of the extended specimens being developed in the State, and implementation of the ES, the scope of ES to include all the biology, geography, and all the images, and everything connected to the specimen
- Extended species is a powerful tool, and it starts with physical specimens, and the primary component is the basic digitisation of the specimens, then in the secondary connection, the linkage will be made, Tertiary extension will be made using eg traits and description
- BCON heading in the same direction
- There is a need in Europe for data infrastructure
- In DiSSCo the idea was the DS provides anchoring function for all kinds of the data form
- DiSSCo keep eyes on European Science cloud to make sure all the data is FAIR
- DS should be the twins for physical specimen
- Where to keep the ES? Some suggestions were in GBIF, but what about the Institution values? Or put in the collection management system, but it might be not satisfying, it is not possible e to keep all the links to the third-party data, and is to hard to adap all the management system to do this
- In DiSSCo it will be done in a unique system with the added values for members and governors
- DiSSCo recognise the small institutions which are unable to manage the ES, a cor infrastructure is needed
- Dissco has a plan for achieve a combination approach, using a collection management system using link tools such as BioCase
- Achieve with a combination approach CMS + ES/DS Store + HUB Service Functions
- The collection management and the ES is the join responsibility of DiSSCo and the institution, and also can be done through a national node
- ES can be provided as a service by DiSSCo
- Upcoming in the next 6 months: specify the need based on CORDRA, how this will be built, and how to mobilise this
- The basic component are applicable for all sort of institution
- ICEDIG has recommended the use of DIgital Object Architecture and CORDRA is the reference implementation for that.

- Talk by Sharif: Gardner Hype Cycle (Ontologies/Graphs are on the rise), link als to the technical fundamentals

Hype Cycle for Emerging Technologies, 2020

Expectations

- Secure Access Service Edge (SASE)
- Social Distancing Technologies
- Embedded AI
- Data Fabric
- Explainable AI
- Composable Enterprise
- AI Augmented Development
- Carbon-Based Transistors
- Responsible AI
- Multiexperience
- Digital Twin of the Person
- Packaged Business Capabilities
- Generative AI
- Composite AI
- Adaptive ML
- Citizen Twin
- Bring Your Own Identity
- Social Data
- Private 5G
- Generative Adversarial Networks
- Differential Privacy
- Biodegradable Sensors
- 2-Way BMI (Brain Machine Interface)
- Ontologies and Graphs
- Health Passport
- Self-Supervised Learning
- Low-Cost Single-Board Computers at the Edge
- DNA Computing and Storage
- AI-Assisted Design
- Authenticated Provenance

Innovation Trigger | Peak of Inflated Expectations | Trough of Disillusionment | Slope of Enlightenment | Plateau of Productivity

Time

Plateau will be reached:
○ less than 2 years   ● 2 to 5 years   ● 5 to 10 years   ▲ more than 10 years   ⊗ obsolete before plateau   As of July 2020

gartner.com/SmarterWithGartner

Source: Gartner
© 2020 Gartner, Inc. and/or its affiliates. All rights reserved. Gartner and Hype Cycle are registered trademarks of Gartner, Inc. and its affiliates in the U.S.

Gartner

Software Development
Architecture and Design 2020 Q2 Graph
http://infoq.link/architecture-trends-2020

InfoQ

Innovators	Early Adopters	Early Majority	Late Majority
Micro frontends	Modular Monolith	GraphQL	Microservices
Data Mesh	Correctly built distributed systems	Reactive programming	Domain-driven Design
AsyncAPI	Actor Model	HTTP/2 and gRPC	Behavior-driven Design
Policy as Code	"Serverless"	Event-driven architecture	Test-driven design
Blockchain	Service Mesh	CQRS	REST
HTTP/3	Functional programming	Event sourcing	
RSocket & Reactive Streams	"Architect as technical leader"	Eventual consistency	
	workflow and decision automation platforms		
	Low code/no code		

CHASM

- From system centric view to information centric view, DO fabric
- DiSSCO Building blocks: Digital Specimen at the bottom, aggregation of the DO on top -> On top of the aggregations build the services/ collections overview Links: Digital Object Interface Protocol specification: https://hdl.handle.net/0.DOIP/DOIPV2.0
- Paper link: http://doi.org/10.5334/dsj-2020-050
-
- Upcoming task: Incorporation of RDA outputs

Discussion from the Chat:
- Alex Hardisty: ICEDIG has recommended the use of DIgital Object Architecture and CORDRA is the reference implementation for that.
- Roger Hyam: Any example deployments?: -> http://dtr-pit.pidconsortium.net/#urls/intro.html (cordra deployment)
- Alex Hardisty: Search for your fave film here: https://ui.eidr.org/search Has been running for 7 years now and powers the movie/TV industry supply chain, Majority of DOI

infrastructure, Majority of DOI infrastructure is powered by CORDRA, UK construction industry and global financial derivatives sector are adopting

- Anton Güntsch: /question I think the Magic box is the 'automated linking' in Alex' diagram. Presently, many collections are setting links to external resources localy in their CMS. Should These activities be moved to a central infrastructure? And how can we synchronise CMS <-> centrally held links <-> external resources?
  Alex: Paid not very much attention to linking so far. Botany pilot (WP5 task?).
- Tim Robertson: "Majority of DOI infrastructure is powered by CORDRA." My understanding it that parts of the Cordra stack (e.g. handles) yes, but not for the majority of the actual objects managed. DataCite for example doesn't use Cordra
- Alex Hardisty: @Tim, you're correct. But CORDRA could be used for that. And we will for the DiSSCo pid scheme. Can read about use of DO architecture in biodiversity and geodiversity here: https://doi.org/10.1162/dint_a_00034
- Roger Hyam: the git repo for CORDRA https://gitlab.com/cnri/cordra/cordra
- Alex Hardisty: @Roger, yes v2.2.0 is the current version. We're running a slightly older version in nsidr.org
- Roger Hyam: @Alex why only one developer on the git? Why no issues? Where is the dev community?
- Alex Hardisty: @Roger, majority of the development is presently done by CNRI Inc. and then made open there. See also https://www.cordra.org/ for the release channel.
- Alex Hardisty: The paper Sharif mentioned explaining this work is: Incorporating RDA Outputs in the Design of a European Research Infrastructure for Natural Science Collections. https://doi.org/10.5334/dsj-2020-050   Someone earlier this week said one of their biggest problems was how to get hold of the original, hi resolution image of an object. What Sharif describes here helps.
- Dimitris Koureas: The first international FAIR Digital Objects conference is going to take place early next year
- Patricia Mergen: For existing EOSC services check here : https://eosc-portal.eu/services-resources
- For quick overview of EOSC read this document : https://op.europa.eu/en/publication-detail/-/publication/581d82a4-2ed6-11eb-b27b-01aa75ed71a1/language-en/format-PDF/source-175468053 https://marketplace.eosc-portal.eu/services/gbif-spain-collections-registry/information
- GBIF SPAIN is for example registered in the EOSC services portal with collection registry and images server.
-

Questions/ Discussion:
- How established the FAIR DO Architecture
- Big transition in the underlying infrastructure towards EOSC
- Do all agree on the importance of the synchronization/ linking … should we stop linking objects locally and move this all to a centrally organized repository? -> Topic of linking has not yet received a huge amount of attention, need to focus more on it (in the coming 6 months). This is an international/global issue. Consider also bidirectional linking, is even more challenging.
- Elspeth Haston points out IIIF work?
- Tim Robertson: Related to the linking. The automated approach to linking (clustering) of records in GBIF can find some(!) links to contribute (e.g. duplicate specimens, sequences

for specimens etc)
https://www.gbif.org/news/4U1dz8LyqQvqIywiRIRpAU/new-data-clustering-feature-aims-to-improve-data-quality-and-reveal-cross-dataset-connections

- Dimitris Koureas: Linking is happening of course already at multiple levels and scales. Is happening at institutional level, individual researcher level, community and aggregator level. Links are hidden in literature, systems and institutional CMSs. AI/ML is helping a lot, but still efforts are fragmented. DiSSCo will need to harness many of these sources of links and allow the community to build upon them in the community curation model it is promising.
- Alex Hardisty: I wrote this blog post on linking in July 2019: https://alexhardisty.wordpress.com/2019/07/25/building-the-dissco-knowledge-graph/ I have a vision for the 'Tahana Link Builder' as a portfolio or box of tools to support all the different kinds of linking, that could be incorporated into workflows. This can develop and grow gradually but it must do so in line with some guidelines that still have to be set.
-

Talk from Mathias Dillen about Biohackathon 2020: Connecting molecular sequences to their voucher specimens:

- General task: Connection of sequences to digital objects, pushing to digital object store
- Conclusions from this Biohackathon: Importance of the incorporation of PIDs into the data model for people, institutions, publications and specimen
- Obtain a PID for a specimen at earliest point possible, train collectors on the correct storage of the data
- repo link: https://github.com/elixir-europe/BioHackathon-projects-2020

Questions/ Discussion:

- Has anyone more uptodate information? -> data structures are very old and difficult to change. But: Situation is beginning to change, e.g., ELIXIR. A white paper is being developed by the ELIXIR Biodiversity Focus Group.
- Jiri Frank: A bit off-topic. ELIXIR is also indeed open to discuss collaboration with DiSSCo. At least from the feedback from the ELIXIR-CZ node director.
- Patricia Mergen: In Belgium we already collaborate with them and other RIs see : https://www.fwo.be/media/1023881/iri_2020_connecting_esfris_frederikcoppens.pdf
- Wouter Addink: @Jiri note also that DiSSCo is going to very closely collaborate with ELIXIR and other RIs in the BiCIKL project
- Mathias Dillen: @Tim: How would you in GBIF link those sequence-based occurrences to their corresponding specimen-based occurrences?
- Tim Robertson: We use names, locations, dates, identifications, typification status, and all local identifiers to "fuzzy" match between records (Think nearest neighbour with thresholding)
- Mathias Dillen: Can data providers contribute to the linking process (provide explicit links or inform on what to look for)?
- And once links are made, how do you model that both the occurrence records are tied to the same occurrence (i.e. gathering event)?
- Tim Robertson: "Can data providers contribute". Not yet Matthias. There are 2 aspects to this - one is using things like "resource relationships" to assert links, and the other is through using external files that help disambiguate things.
  At this stage, we've refrained from asserting definitive links. We present them as "hints" of

related records with justification as to why they are being linked. If we get to the stage of being sure the links are correct, effectively each record can participate in the cluster, and all those records provide some evidence of the occurrence.
- Elspeth Haston: @Tim and @Mathias. I would be keen to share what we looked at some time ago on linking vouchers to specimens. I think it potentially also falls into the work of SYNTHESYS NA3 & JRA2
- Tim Robertson: Just to be clear: GBIF's work here is infancy and there is a lot that could be (needs to be) done to evolve this.
- Eslpeth Haston: Getting better data into the molecular data records will help GBIF and everyone in the future.
- Wouter Addink: supporting data providers in using globally unique identifiers for agents, specimens and other data will certainly help
- We plan to publish a link to this morning session recordings through a news item soon
-
-

Discussion from the chat:
- Anton Güntsch: Standardising specimen citations in sequence data will be a task in the BiCIKL project. This will definitely help finding them.
- Laurence Livermore: @Anton is there a link somewhere for the BiCIKL project?
- Anton Güntsch: @Laurence: BiCIKL will start in May. I don't think there is already a web presence.
- Tim Robertson: GBIF currently only have the geo referenced records from EMBL (1.4M specimens) in the GBIF index. We're working with EMBL to prepare a dataset with all records from EMBL relating to specimens (8M records). This will help with what Matthias is presenting, but (sorry) does mean some of it may need redone. You can explore the EMBL dataset on the GBIF UAT environment here https://www.gbif-uat.org/occurrence/search?basis_of_record=PRESERVED_SPECIMEN&dataset_key=ad43e954-dd79-4986-ae34-9ccdbd8bf568
- Alex Hardisty: Mathias: great that you described the two different ways (options) of representing a link in a DS.
-
-

Talk by Claus Weiland on: SEMAF
- SEMAF focuses on ontologies. Aim: To assess existing ontologies from different domains and how can improved, shared, …
- https://www.eoscsecretariat.eu/Supporting-Interoperability-EOSC-Flexible-Semantic-Mapping-Framework

Questions/ Discussion:
- Remarks: Important topic because it makes data not only machine readable but also machine accessible.

Talk by Alex Hardisty on OpenDS and on MIDS:
- Positioning openDS in the landscape:
  https://github.com/DiSSCo/openDS/blob/master/positioning-opends.md To follow the opends work: https://github.com/DiSSCo/openDS ES/DS global consulation, blog post on website: https://www.allianceforbio.org and background document:
  http://bit.ly/esdsconsult
- NSF funded US prototype
- Developed stronger bonds between DiSSCo/iDigBio,

Questions/ Discussion:
- How safe is the funding/ grant? We don't know yet.
- Is it feasible for us to do further development for digital object management? E.g. EOSC proposal? -> This could be a possibility if we find appropriate partners (like ELIXIR).
- EOSC: option for smaller calls, Pat: Align with CETAF
- When should we start discussing how CMS should synchronize with digital object stores?
  -> Best: Straight away! Start the discussion in the meeting next monday.
  -> Tim: Data mapping critical

Saved Chat

09:00:10   From Claus Weiland :
https://docs.google.com/document/d/1XD3EyqGbISlZmfy9yJJRyUTQx7gwBupg0Uio9Rusr6c/edit
?usp=sharing
09:05:15   From Jiri Frank :
https://docs.google.com/document/d/1XD3EyqGbISlZmfy9yJJRyUTQx7gwBupg0Uio9Rusr6c/edit
?usp=sharing
09:09:49   From Claus Weiland :
https://docs.google.com/document/d/1XD3EyqGbISlZmfy9yJJRyUTQx7gwBupg0Uio9Rusr6c/edit
Claus Weiland :
https://docs.google.com/document/d/1XD3EyqGbISlZmfy9yJJRyUTQx7gwBupg0Uio9Rusr6c/edit
#
09:30:52   From Alex Hardisty : Current working draft of MIDS here:
https://github.com/tdwg/mids/blob/working-draft/current-draft/MIDS-definition-v0.12-03Nov2020
.md
09:31:26   From Alex Hardisty : If you want to contribute to the TDWG Task Group you can sign up to its mailing list here: http://lists.tdwg.org/mailman/admin/tdwg-mids
09:35:38   From Ana Casino : Alex, please, confirm the access to the mailman list, since it requests the admin login password
09:36:13   From David Fichtmueller : http://lists.tdwg.org/mailman/listinfo/tdwg-mids
09:36:26   From David Fichtmueller : this is the link for subscribing
09:36:38   From Ana Casino : Thanks David
09:37:23   From Pieter Huybrechts (MeiseBG) : +1 Elspeth, I stress that the capture of missing data is very valuable
09:40:40   From Alex Hardisty : Sorry @Ana, I got the wrong link. @David pasted the correct one. Here it is again for signing up to the MIDS mailing list:
http://lists.tdwg.org/mailman/listinfo/tdwg-mids

09:41:43   From Wouter Addink : There is a DwC field for institution referent, which is dwc:institutionID. Issue with that field though is that it includes a recommendation to use an identifier from a collections registry, and cannot capture the type of identifier (GRID, ROR etc)

09:46:04   From Niels Raes : TDWG suggests http://biocol.org/urn:lsid:biocol.org:col:34777. Why cannot https://ror.org/04py0zz23 be used?

09:46:54   From Mathias Dillen : I think because the documentation is quite old

09:49:25   From Wouter Addink : @Niels it could be used. It is just against the current recommendation.

09:51:36   From Mathias Dillen : Even if we don't have institutionIDType, the use of institutionID will resolve quite a bit of ambiguity already compared to institutionCode.

09:52:19   From Alex Hardisty : https://github.com/DINA-Web

09:52:43   From Alex Hardisty : https://www.dina-project.net/wiki/Welcome_to_DINA

09:57:58   From Patrick Semal : The risk is the same with a commercial company ...

10:00:02   From Josh Humphries : a point on the security of open source projects - the benefit of security issues being spotted in open source projects is only relevant if there are lots of eyes looking at the code.

10:08:47   From Alex Hardisty : WP6 technical background presentations: http://bit.ly/DiSSCoWP6Share Explanatory video, 12 minutes: https://bit.ly/esdsframework Positioning openDS in the landscape: https://github.com/DiSSCo/openDS/blob/master/positioning-opends.md To follow the opends work: https://github.com/DiSSCo/openDS

10:29:46   From Roger Hyam : Why Cordra? When and how was that decision made?

10:32:29   From Wouter Addink : @Roger in the ICEDIG project

10:32:30   From Alex Hardisty : @Roger. ICEDIG has recommendated the use of DIgital Object Architecture and CORDRA is the reference implementation for that.

10:33:19   From Roger Hyam : Any example deployments?

10:37:00   From Sharif Islam : http://dtr-pit.pidconsortium.net/#urls/intro.html (corder deployment)

10:37:05   From Sharif Islam : cordra

10:41:54   From Alex Hardisty : @Roger: Search for your fave film here: https://ui.eidr.org/search Has been running for 7 years now and powers the movie/TV industry supply chain.

10:42:18   From Alex Hardisty : @Roger: Majority of DOI infrastructure.

10:42:45   From Alex Hardisty : @Roger: Majority of DOI infrastructure is powered by CORDRA.

10:44:36   From Alex Hardisty : @Roger: UK construction industry and global financial derivatives sector are adopting.

10:47:47   From Anton Güntsch : /question I think the Magic box is the 'automated linking' in Alex' diagram. Presently, many collections are setting links to external resources localy in their CMS. Should These activities be moved to a central infrastructure? And how can we synchronise CMS <-> centrally held links <-> external resources?

10:48:04   From Tim Robertson (GBIF) : "Majority of DOI infrastructure is powered by CORDRA." My understanding it that parts of the Cordra stack (e.g. handles) yes, but not for the majority of the actual objects managed. DataCite for example doesn't use Cordra

10:49:15   From Alex Hardisty : @Tim, you're correct. But CORDRA could be used for that. And we will for the DiSSCo pid scheme.

10:53:39   From Alex Hardisty : Can read about use of DO architecture in biodiversity and geodiversity here: https://doi.org10.1162/dint_a_00034

10:54:07   From Roger Hyam : @Alex Is this the git repo for CORDRA
https://gitlab.com/cnri/cordra/cordra
10:54:45   From Wouter Addink : link had a typo: https://doi.org/10.1162/dint_a_00034
10:54:54   From Alex Hardisty : @Roger, yes v2.2.0 is the current version. We're running a slightly older version in nsidr.org
10:55:42   From Roger Hyam : @Alex why only one developer on the git? Why no issues? Where is the dev community?
10:56:46   From Alex Hardisty : @Roger, majority of the development is presently done by CNRI Inc. and then made open there. See also https://www.cordra.org/ for the release channel.
10:58:25   From Roger Hyam : I have another meeting. Hope to come back later.
10:59:51   From Alex Hardisty : The paper Sharif mentioned explaining this work is: Incorporating RDA Outputs in the Design of a European Research Infrastructure for Natural Science Collections. https://doi.org/10.5334/dsj-2020-050
11:03:50   From Alex Hardisty : Someone earlier this week said one of their biggest problems was how to get hold of the original, hi resolution image of an object. What Sharif describes here helps.
11:07:31   From Dimitris Koureas : The first international FAIR Digital Objects conference is going to take place early next year
11:07:58   From Patricia Mergen : For existing EOSC services check here : https://eosc-portal.eu/services-resources
11:10:56   From Patricia Mergen : For quick overview of EOSC read this document : https://op.europa.eu/en/publication-detail/-/publication/581d82a4-2ed6-11eb-b27b-01aa75ed71a1/language-en/format-PDF/source-175468053
11:18:17   From Patricia Mergen : https://marketplace.eosc-portal.eu/services/gbif-spain-collections-registry/information
11:18:23   From Elspeth Haston : IIIF work?
11:18:43   From Tim Robertson (GBIF) : Related to the linking. The automated approach to linking (clustering) of records in GBIF can find some(!) links to contribute (e.g. duplicate specimens, sequences for specimens etc)
https://www.gbif.org/news/4U1dz8LygQvqIywiRIRpAU/new-data-clustering-feature-aims-to-improve-data-quality-and-reveal-cross-dataset-connections
11:19:47   From Dimitris Koureas : @Anton: Linking is happening of course already at multiple levels and scales. Is happening at institutional level, individual researcher level, community and aggregator level. Links are hidden in literature, systems and institutional CMSs. AI/ML is helping a lot, but still efforts are fragmented. DiSSCo will need to harness many of these sources of links and allow the community to build upon them in the community curation model it is promising.
11:21:48   From Alex Hardisty : I wrote this blog post on linking in July 2019: https://alexhardisty.wordpress.com/2019/07/25/building-the-dissco-knowledge-graph/
11:24:21   From Alex Hardisty : I have a vision for the 'Tahana Link Builder' as a portfolio or box of tools to support all the different kinds of linking, that could be incorporated into workflows. This can develop and grow gradually but it must do so in line with some guidelines that still have to be set.
11:27:46   From Anton Güntsch : Standardising specimen citations in sequence data will be a task in the BiCIKL project. This will definitely help finding them.
11:28:47   From Laurence Livermore : @Anton is there a link somewhere for the BiCIKL project?
11:30:27   From Anton Güntsch : @Laurence: BiCIKL will start in May. I don't think there is already a web presence.

11:32:47   From Tim Robertson (GBIF) : GBIF currently only have the geo referenced records from EMBL (1.4M specimens) in the GBIF index. We're working with EMBL to prepare a dataset with all records from EMBL relating to specimens (8M records). This will help with what Matthias is presenting, but (sorry) does mean some of it may need redone. You can explore the EMBL dataset on the GBIF UAT environment here
https://www.gbif-uat.org/occurrence/search?basis_of_record=PRESERVED_SPECIMEN&dataset_key=ad43e954-dd79-4986-ae34-9ccdbd8bf568
11:33:06   From Alex Hardisty : @Mathias: great that you described the two different ways (options) of representing a link in a DS.
11:36:03   From Josh Humphries : repo link:
https://github.com/elixir-europe/BioHackathon-projects-2020
11:40:27   From Niels Raes : A white paper is being developed by the ELIXIR Biodiversity Focus Group
11:46:25   From Jiri Frank : A bit off-topic. ELIXIR is also indeed open to discus collaboration with DiSSCo. At least from the feedback from the ELIXIR-CZ node director.
11:47:02   From Alex Hardisty : @Jiri, yes, I believe that is true.
11:48:10   From Patricia Mergen : @Jiri In Belgium we already collaborate with them and other RIs see : https://www.fwo.be/media/1023881/iri_2020_connecting_esfris_frederikcoppens.pdf
11:48:30   From Jiri Frank : Excellent, we are starting this year to use their data services.
11:55:31   From Wouter Addink : @Jiri note also that DiSSCo is going to very closely collaborate with ELIXIR and other RIs in the BiCIKL project
11:57:28   From Mathias Dillen : @Tim: How would you in GBIF link those sequence-based occurrences to their corresponding specimen-based occurrences?
11:58:45   From Tim Robertson (GBIF) : We use names, locations, dates, identifications, typification status, and all local identifiers to "fuzzy" match between records
11:59:11   From Tim Robertson (GBIF) : (Think nearest neighbour with thresholding)
11:59:51   From Alex Hardisty : Positioning openDS in the landscape:
https://github.com/DiSSCo/openDS/blob/master/positioning-opends.md To follow the opends work: https://github.com/DiSSCo/openDS ES/DS global consulation, blog post on website: https://www.allianceforbio.org and background document: http://bit.ly/esdsconsult
12:03:52   From Mathias Dillen : Can data providers contribute to the linking process (provide explicit links or inform on what to look for)?
12:04:43   From Mathias Dillen : And once links are made, how do you model that both the occurrence records are tied to the same occurrence (i.e. gathering event)?
12:05:46   From Tim Robertson (GBIF) : "Can data providers contribute". Not yet Matthias. There are 2 aspects to this - one is using things like "resource relationships" to assert links, and the other is through using external files that help disambiguate things.
12:07:19   From Tim Robertson (GBIF) : At this stage, we've refrained from asserting definitive links. We present them as "hints" of related records with justification as to why they are being linked. If we get to the stage of being sure the links are correct, effectively each record can participate in the cluster, and all those records provide some evidence of the occurrence.
12:07:27   From Elspeth Haston : @Tim and @Mathias. I would be keen to share what we looked at some time ago on linking vouchers to specimens. I think it potentially also falls into the work of SYNTHESYS NA3 & JRA2
12:08:34   From Tim Robertson (GBIF) : Just to be clear: GBIF's work here is infancy and there is a lot that could be (needs to be) done to evolve this.
12:08:49   From Elspeth Haston : Getting better data into the molecular data records will help GBIF and everyone in the future.

12:09:55   From Wouter Addink : supporting data providers in using globally unique identifiers for agents, specimens and other data will certainly help
12:13:24   From Wouter Addink : We plan to publish a link to this morning session recordings through a news item soon

DPP All Hands Meeting 1

T7.1 Governance structure – Core elements and landscape analysis - Main outcomes

Eva M. Alonso.

DiSSCo Prepare Project Manager

Task 7.1 Leader – Refinement of the governance model, strategy & operational planning

Funded by the Horizon 2020
Framework Programme
of the European Union

Objective: Refinement of the interim DiSSCo governance model
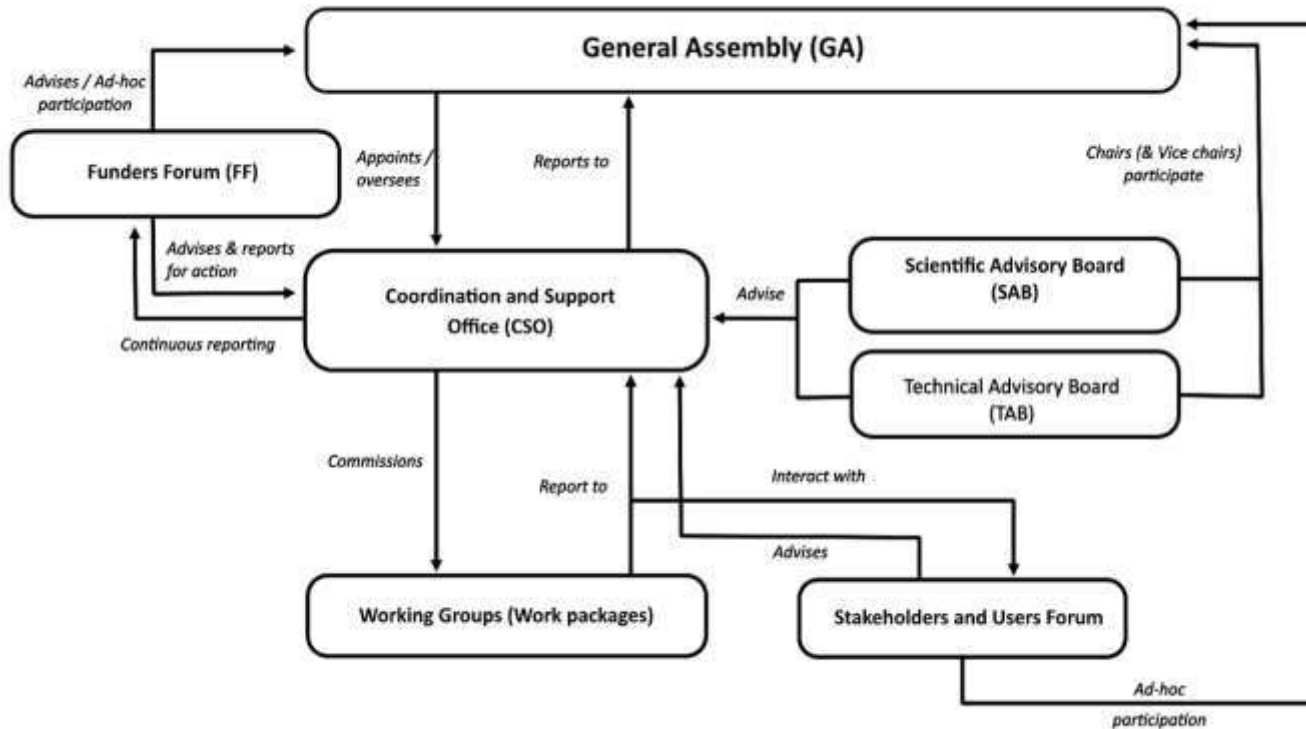
That implies
- to be instrumental to the DiSSCo LE
- + Strategy and strategic planning

Purpose of the brainstorming session

✓ Identify key questions and core elements in the governance structure proposed,
✓ Agreed on methodology and timeline.

Highlights

1. Almost 40 participants. Extremely good contributions from Beneficiary partners outside the original team.
2. Alignment with T7.2 LE better articulated now
3. Methodology agreed upon:
   a. Bi-monthly meetings > Identification of core governance elements + analysis of best practices (incl.ERIC Forum),
   b. Consultation processes > Work timeline align with DiSSCo meetings calendar (NNs, FF, GA),
   c. Coordination with DiSSCo Aspiration WG to refine DiSSCo mission/vision > DiSSCo strategy.

DiSSCo All Hands - Task 7.1

DiSSCo Governance Structure - Core elements & Landscape analysis

Monday 18th January 2021, 15.00-16.30 CET

Notetaker(s): Tina Loo & Marie-Laure Kamatali

Participants:
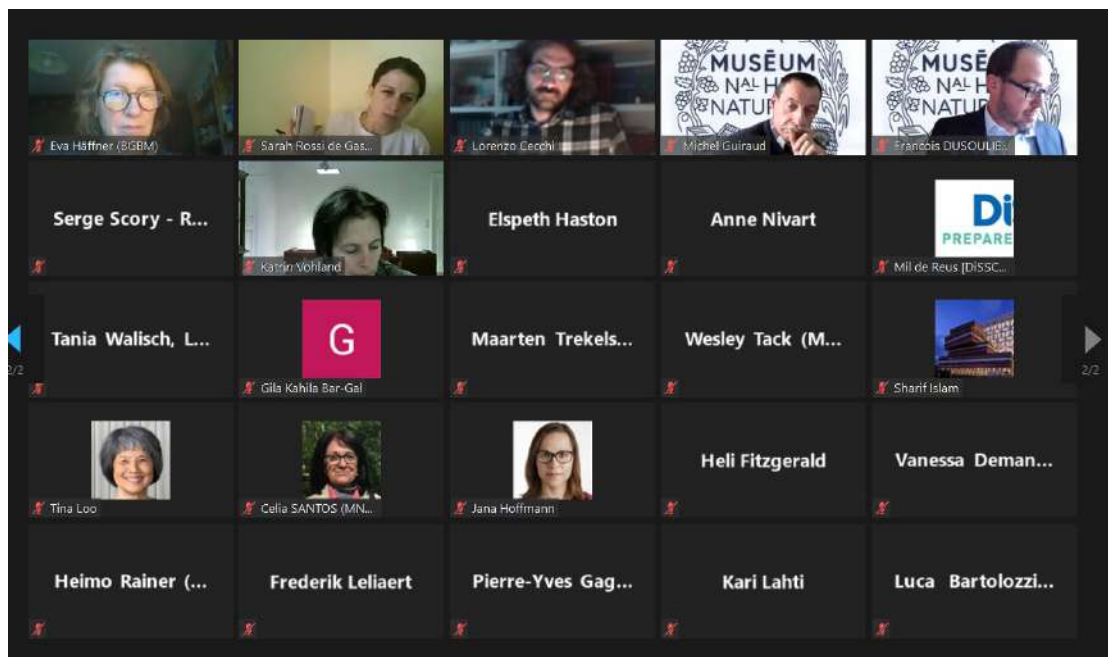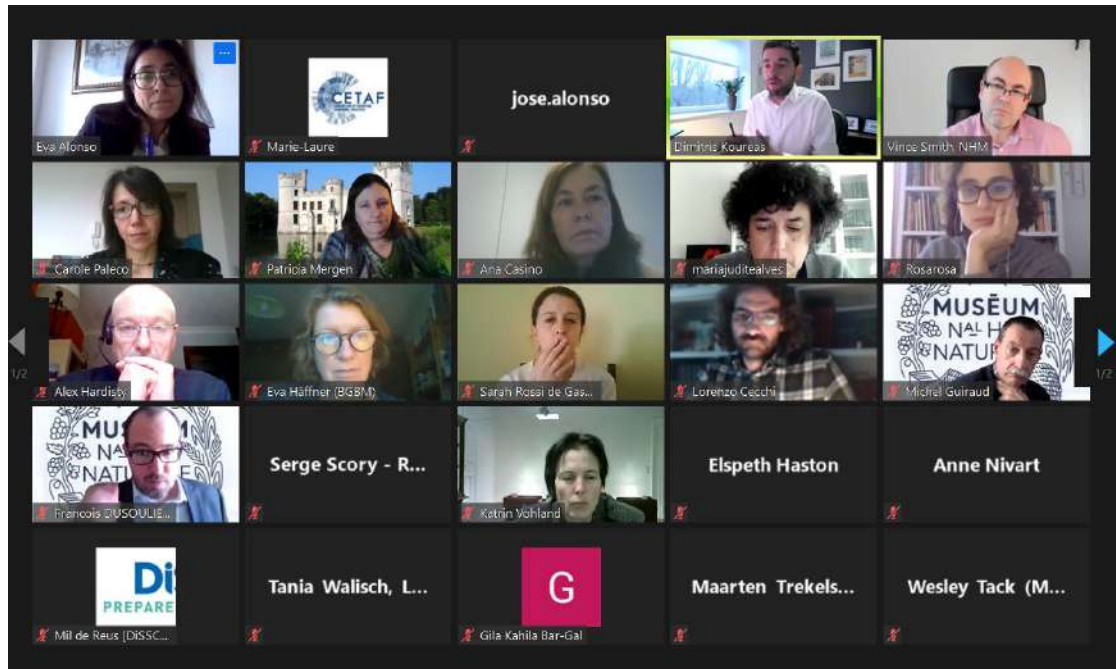
Purpose of session

This session aims to identify core elements agreed upon by the Consortia and new ones to include in the governance structure proposed. Secondly the discussion aims to set up requirements to analyse existing governance models in other RIs that facilitates the definition of a governance structure and its operation.

Both elements will contribute to the preparation of MS7.4 update/upgrade of the governance model that will respond to the needs of the legal entity to be in place.
The work in T7.1 keeps an strong dependency with T7.2 which currently is about to deliver a proposal on the future legal entity.

Expected outcomes

The session will deliver
- List of core elements of the governance chart,
- List of requirements and RIs (best practices)
- Agreement on the timeline scheduled

Supporting documentation

Agenda of the session

15:00 - 15.10	Welcome and introductions
15:10 - 15:20	Presentation of the rationale and initial work done in T7.1 and supporting documentation.
15:20 - 15:50	Discussion of the core elements - Outcome: List of core elements, existing or new, that need to be included in the governance model (MS7.4)
15:50 - 16:10	Discussion of the requirements to identify best practices - What are the essential requirements a RIs must have to be considered a best practice. - Outcome: List of RIs to analyse in-depth
16:10 - 16:25	Discussion of the next steps - Structure of the MS.7.4 - Consultation process - Timeline for MS7.3 Strategy and operational planning
16:25 - 16.30	AoB

Notes

Initial outcomes:

- DiSSCo EU MoU: important document for DiSSCo that comprises two annexes including governance model across different stages of DiSSCo. CUrrently in the signing process.
- Existing interim gov chart and core elements in the actual description of the interim gov model. Needs to be discussed but already in place
- Need to refine DiSSCo mission and vision: as part of the work around the strategy of DiSSCo.
- Close cooperation with T7.2 and T8.1 as NN consultation and consus are highly important to move task forward

Main objective: the governance model schema which needs to be continuously worked on. Currently there are two modes for LE (ERIC and AISBL) which both fit this gov model.
Question: does this model fulfill our vision for DiSSCo? Not really
To move forward; need to ask ourselves right questions

Starting point: DISSCo specificities
Need to identify key important questions:

1) What are the core elements that we need to preserve? DiSSCo has already started to collect and define what core elements will be part of the RI both technical and organisational
2) Which provisions will guarantee long term sustainability (financially and service provision)
3) How will institutions participate meaningfully in the decision making process and in the review/upgrade of services and policies? How will they be represented?
4) What will be the contribution model? How in-kind contributions will be articulated to be meaningful and visible?
5) What kind of governance do we want? Flat or what degree of hierarchy?

Many questions that try to address the specificities of DiSSCo and most likely more questions.

Alex: Additional questions:

- How will DISSCo coordinate its international relations? (peer infrastructure colleagues globally, GBIF, Lifewatch, Elixir,etc) How will this relationship be governed?
- How will DiSSCo be governing the services it commits to providing?

Vince:

- international dimension and relationship with other international RIs
- Relationship with institutions we want and the type of engagement we want. Institutions will be continuously engaged and they will be driving this relationship and how is that best preserve in the model going forward?

Jose: How are we going to manage to address and coordinate the various cultural visions of DiSSCo? Work to come up with a flexible approach so that everyone feels included and that they are contributing. (EA: geographical equity is always a concern)

Alex: The cultural visions comment goes with the comment regarding preserving ongoing relationships with institutions. Approaches vary from task focused to relationship focused. Mechanism we opt for has to appeal to both technical and human relations aspects.

Starting point: look for Best Practices - Landscape analysis: ERIC/AISBL
Frequent question: How can we consider one RI as BP?
Identify five or six requirements shared by DiSSCo and other RIs; we go beyond our domain. Look at how other ERICs and AIBLs are doing in terms of governance, what are the benefits and challenges of specific gov models.
Requirements:
1) Look for country based RIs
2) Look for community rooted RIs: identification of needs shared by huge amount of researchers, for e.g.
3) Distributed data-driven RI:

Eva clarifies: want to analyze other existing RIs, identify series of criteria that go behind choosing a specific model. Based on these criteria, select 5 RIs for which we will analyze the governance model.

Comments:
- MG - governance reflects the activity, the purpose of the RI. Therefore, suggests to look at RIs that have the same purpose as DiSSCo (similar purpose criterion).
- DK: there are two things that are linked but to be distinguished. First, focus on defining requirements (governance traits that need to be included to allow the community to feel at ease) and second (which RI has a governance model that respects one of or all requirements) and suggests a modular approach.

This exercise might be premature as it assumes that DiSSCo will opt for the ERIC (consultation still needs to take place. The legal advisor recommends opting for ERIC, as mentioned in the report; from the beginning, it was established that DiSSCo will be country based making it less relevant to consider AIBSL.

Vince - still has questions about AIBSL and ERIC as the report still mentions that ERIC and AISBL are both viable, and need to discuss the report.

Patricia shares EOSC insights that opted for AISBL due to tight timeline but preferred choice on the long term is an ERIC.

Ana as a way to overcome this challenge: instead of going directly to the landscape analysis, perhaps we could start with which requirements that should be in place. Do we have advisory bodies? How will we connect with the scientific community? Will they have a say in the decision making process? Then status, bylaws will follow. Which are the landmarks already operational that meet all the defined essential requirements for RIs.

Serge - there is at least one AISBL (not an RI) where states are the members represented by members of agencies not ministries. But it costs. In the report from the legal advisor, Serge suggests to go review the report and will provide a slide focusing on costs.

Gildas comments that Paris is more in favor of country based membership (like ERIC) as it is a very centralized country and also in terms of sustainability, ERIC is the best outcome. Furthermore, DiSSCo is very technical, data focused RI: need to take into consideration,as a requirement, that the governance model facilitates the communication between technical workers and policy workers.

Dimitris: forward independent of the legal entity, we have a set of questions regarding participation but by the end of the exercise what we need to see is how those requirements of organisations of stakeholders that will be active in the management of the RI. They will need to translate in organisational bodies There are different ways of combining them, different levels of granularity. Importance of understanding the following: international level, role of the bodies, scope of their activities. How the legal entity will affect that is to see but it is currently not a dealbreaker.

Agenda item 3 Timeline

Different milestones and deliverables expected at the end of the project.
April 2021 consultation process during NN meeting - connection with T8.1
May 2021 FF consultation
June 2021 GA: think about whether we can present elements of our work
Currently planning to have two meetings per month; some only in the presence of the task force (includes only the partners with most PMs (MNHN and RBINS need to be committed but open to all). CETAF is included as they are part of CSO, DiSSCO general management. Strong dependencies with NN and FF meetings.

Vince Smith: the role of the FF, how will it be connected to all of this? Timeline sounds fine.

Eva: FF inaugural meeting is 25 February 2021 but will not present anything. The FF is an advisory body, highly political with governmental representatives. They will have a say and their insights will help us align with their expectations. Importance of having it in place from day one.

Dimitris: better to keep countries engaged and informed since they will be funding eventually. Mostly on organisational, financial, government aspects are most interesting to them.

Next steps: align with T7.2 LE need for clear understanding of how we will
Whatever LE will be our model, if we can start working on the implementation of this entity model in 2022 or early 2023, that would be the best.

Open discussion

Vince: to watch closely is the relationship between FF and GA and how that may evolve moving forward.

Dimitris: it is important to not limit ourselves to the terminology we currently use in the context of the interim governance. It is a starting point but not a definitive roadmap. If we decide to go towards ERIC, the FF will be the ultimate decision making body. But FF will disappear by the end of this process.

Four major takeaways from this conversation:
- Major topic: regarding the operation of DISSCo which relies almost entirely on the contribution of the institutions. At what level of organisation we see institutions participate in the decision making process of DiSSCo? What model fits? Through CETAF? Through a different model that includes all the models from institutions partners? What challenges are foreseen?
- How people with technical knowledge drive the RI, which is a common trait for RIs. How do we reconcile that the RI is data driven but in its essence a community initiative? What is the proportionality based on contributions (in-kind and monetary) of each country?

Michel - again back to the purpose of DiSSCo. To summarize it there to get information out of the collections for digitization and the wider scientific community. Four pillars: funds, scientific forum, collection holders and the CSO. Our community has many providers.

Jana Hoffmann shared a wish: DiSSCo to be as open for possibilities as possible. There are constraints due to funding, a must have. But Would like to explore other avenues of contributions and what contributions mean. We could envisage mixed models, not only through state membership. It could be a good point to observe for any future scalability. How we can include other ways, outside of EU formality that create constraints at national level like in Germany. Instead of one or the other, think of mixed models.

Eva reminds that in 2023 DiSSCo faces a gap in the funding program. An early legal entity would facilitate a smooth transition to a new funding model and avoid the risks associated to. That time constraint needs to be taken into account as well.

Vince - Agree with Jana. If the CSO is the engine, the FF is the fuel and the driver, that doesn't leave much room for institutions who are really driving this. There is a risk for institutions to be cut out of the process by the new decision-making body. Mixed model approach that requires a bit more creativity otherwise a lot of time will be spent telling governments what they need to tell DiSSCo (similar to what is done in GBIF).

To close the session Dimitris agrees on the fact need to avoid the risk mentioned by Eva. There is another risk: institutions not coming to this from the same level of maturity and size. Careful not falling into a prevalence model led by the bigger institutions. This was the point of delegating the decision making to countries instead of facilities. How can we first bring in funding from national governments? Elements that we need to balance and find the best way to proceed.

AHM 1

DiSSCo Prepare WP7 – Session T7.2
Presentation of the study on the Legal Entity model(s) proposed and discussion with Legal Advisor as guest

January 20th 2021

Royal Belgian Institute of Natural Sciences

DiSSCo Prepare WP7 T7.2

Session on the "Towards a Legal Model for DiSSCo"

AHM 1 Session Wrap up

Outputs from Session T7.2

o 36 attendees

o Including 15 non WP7 partners

Poll results

☰	Have you been involved in the creation of the legal entity of a Research Infrastructure before?	24 >
💬	If yes which RI (Acronym) ?	10
💬	What is your take home message ?	4

Have you been involved in the creation of the legal entity of a Research Infrastructure before?

Yes
33%

No
67%

Outputs from Session T7.2

Outputs from Session T7.2

o Speakers : partners & Legal Advisor from xOfficio

o Understood that ERIC model is recommended and will be an asset as clear leverage for governmental funding and support

o Need to find the most suited model to allow institutions and CETAF to keep playing a major role

o LE is a complex issue:

  o The model should be made clear to institutions how the decisions to be taken will influence the daily work, activities

  o Further discussions must take place at the various levels (WP8) – bet. Institutions, bet. nodes, bet institutions & governments repres. and between CETAF members

o Next steps : Decision expected from the FF (May) & GA (June) important input for the Deliverable defining the Legal entity model By laws & statutes by Dec 2021

"Bringing the irreplaceable data stored in natural science

collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

The preparatory phase project of DiSSCo
Research Infrastructure - Distributed
System of Scientific Collections

https://www.dissco.eu/prepare/

ALL HANDS MEETING
Minutes of Meeting

DAY 3 Wednesday January 20th 2021

Session 15:00 – 16:30

DiSSCo Prepare WP7 – Session T7.2

Presentation of the study on the Legal Entity model(s) proposed and discussion with

Legal Advisor as guest

Programme of the session

Time	Description	Speaker
15h00 – 15h05	Welcome & Presentation of the session	RBINS CP
15H05 – 15h15	CETAF key role in DiSSCo RI	CETAF, AC
15h15 – 15h25	EOSC RI as an AISBL	MBG, PM
15h25 - 15h35	EMBRC ERIC model and Relations with the French government	MNHN, VD, FD
15h45 – 15h55	Experience with ERIC Forum	Naturalis, EA
15h55 – 16h05	Analysis of the LE models and recommendation for DiSSCo	Ohad Graber-Soudry (Xofficio)
16h05-16h15	Overview of the work and process leading to the Milestone of T7.2	RBINS, SS
16h15 – 16h25	Q&A on key aspects and feedback from participants	All
16h25 – 16h30	Wrap up & Conclusions	RBINS, CP, SS

Convenors : Carole Paleco (CP), Serge Scory (SS)

Notes of the session : Note takers Vanessa Demanoff, François Dusoulier

List of participants:

Name	Institution
Serge Scory	RBINS
Carole Paleco	RBINS
Eva Alonso (EA)	Naturalis
Dimitris Koureas	Naturalis
Tina Loo	Naturalis
Wouter Addink	Naturalis
Ohad Graber-Soudry (OG)	X-Officio
Aino Juslen	LUOMUS
Alex Hardisty	Cardiff University
Ana Casino (AC)	CETAF
Celia Santos	CSIC MNCN
Claus Weiland	Senckenberg
Eva Häffner	BGBM
Falko Glöckler	MfN
François Dusoulier (FD)	MNHN
Gergely Babocsay	NHMH
Gillien	
Heimo Reiner	NHM Vienna
Henrik Enghoff	University Copenhaguen
Jiri Frank	NM CZ
Jose Alonso	Naturalis
Karin Volhand	NHM Vienna
Maria Joao Fonseca	MHNC UP
Marie Judite Alves	ULisboa
Marie Laure Kamatali	CETAF
Michel Guiraud	MNHN
Patricia Mergen (PM)	Meise BG
Patrick Semal	RBINS
Pedro Arsenio	ULisboa
Peter Warth	SMNS
Pierre-Yves Gagnier	MNHN
Quentin Groom	Meise BG
Salome Landel	MNHN
Sara Rossi de Gasperis	Fi NHM
Vanessa Demanoff (VD)	MNHN
Vince Smith	NHM

Minutes
1) The session has been attended by 36 participants. It lasted more thatn the attributed 1h30 as the participants were involved in the end of session discussion regarding the role of the institutions within the recommended model if an ERIC.
2) The meeting started with the presentation by CP of the programme.
   Some questions were asked to the participants through sli-do platform in order to know more about the participants' experience in the creation of a legal entity for a RI.

A third of the participants had an experience with the creation of a LE. From he second question asking in which specific RI they had had an experience with the following RIs were mentioned :

EOSC, REcolnat, Lifewatch, SeaDataNEt AISBL, Dariah. Some mentioned CETAF and TDWG which however are not RIs. This may be understood either as a misunderstanding of the question or lack of information concerning RI and their structure and the framework into which they operate.

See questions and answers below:

≔	Have you been involved in the creation of the legal entity of a Research Infrastructure before?	24
💬	If yes which RI (Acronym) ?	10 >
💬	What is your take home message ?	4

If yes which RI (Acronym) ?

DiSSCo

EOSC

RECOLNAT

TDWG, GBIF

LIFEWATCH

SeaDataNet AISBL

No

LifeWatch and the Extreme Light Infrastructure (ELI)

LifeWatch, DARIAH , CETAF ...
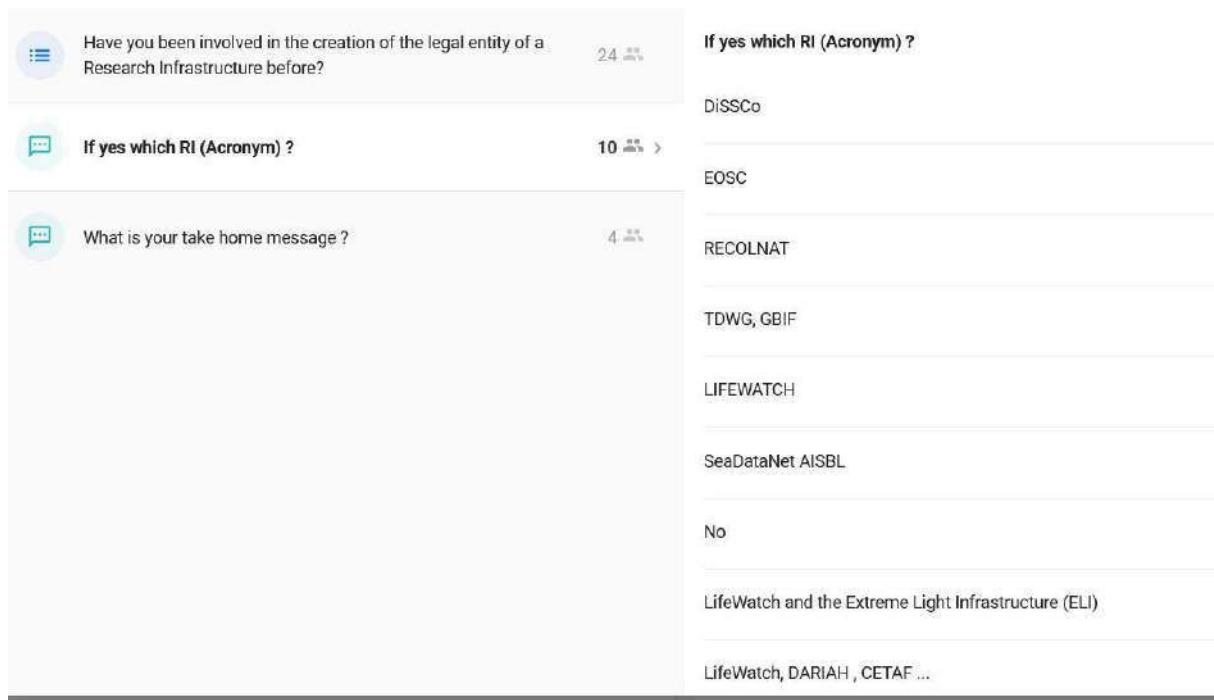
- o SS presented the overall work done towards the MS delivered on 15 January 2021.

- o OG as invited guest from xOfficio presented his work and recommendations for the LE model for DiSSCo and gave the relevant arguments for the ERIC model

  - o "Participation in European Research Infrastructure Consortium (ERIC) should not be a problem as it is open to any third country (and the UK already suggested they will accept the jurisdiction of the Court of Justice). Hosting an ERIC is a different question... "

  - o Some specific concerns dealt with non-EU members joining the RI as ERIC

  - o "In my experience the UK usually have certain requirements that are connected to their status as non-EU member, but they are not the only non-EU members. Israel and Serbia joined a couple of ERICs as well. The UK basically accepted the jurisdiction of the OJEU in relation to disputes between the members and the ERIC or between members. It is a UK negotiation strategy but there is no legal barrier for them to join;"

- o The partners presented their own experience in the field and each gave also recommendations, EMBRC RI and its choice for an ERIC, EOSC choice for an AISBL made for a question of calendar and leaving the Governments not entirely in the decision process. EA presented the advantages of taking part to the ERIC forum for networking, funding opportunities and liaising with other RIs.

DiSSCo
PREPARE

o   During the discussion the participants shared their opinion and notably the fact that they understood that the ERIC model was recommended as an asset and a clear leverage for governmental funding and support.

o   They however stressed the need to find the most suited model to allow institutions and CETAF to keep playing a major role. The role of other organisation could also be considered.

o   In conclusion to the discussion. The choice of a LE is a complex issue and in order to make sure that a large majority of members are reassured :

   o   The model should be made clear to institutions on how the decisions to be taken will influence the daily work, activities

   o   Further discussions must take place at the various levels (WP8) – bet. Institutions, bet. nodes, bet institutions & governments repres. and between CETAF members

o   Next steps and activities to take place:

   o   The decision on the LE model is expected to be raised from the FF (May) & GA (June). In the mean time the T7.2 partners need to address the concerns from the partners and DiSSCo members though consultations notably collecting the impact of other RI ERIC model on their respective members.

   o   The T7.2 partners need to make sure through DPP consultations that the partners abide and approve the model as it will feed the building up of the Deliverable defining the Legal entity model By laws & statutes by Dec 2021.

Outcomes & actions from TASK 7.3 meeting "Develop and establish DiSSCo policies*"
Friday, January 22nd 2021

Leads: Vince Smith & Matt Woodburn (NHM London)
Included 30+ participants from across DPP

January 22nd 2021

1

*not about writing institutional policy or DiSSCo policy, it's a framework for policy alignment & interaction

Focused on user journeys for policy tool

DiSSCo Knowledgebase

Institutional Policies

(policies classified/scored via Synth+ metadata schema)

(Institutional metrics)

Institution dashboard
(maturity index, gaps, comparisons to peer average, more?)

DiSSCo dashboard
(member institutions, listed, ranked and scored)

(services classified/scored via Synth+ metadata schema)

DiSSCo Services

Institution process / view

CSO process / view

"An online checklist tool allowing DiSSCo Partners to map their institutional policies against the DiSSCo Service policy needs to show policy alignment, and for the DiSSCo CSO to see the state of policy compliance across all DiSSCo Partners."

Task 7.3: DiSSCo Policies: outcomes and next steps

- Recognised the complex policy landscape (many policies incomplete for both institutions and DiSSCo Services)
    - *Agreed that a self assessment framework was the right approach*
- Discussed alignment with SYNTH+ 2.1 (policy metadata schema)
    - *Misalignment of timelines can be managed*
- Discussed points of integration with WP5 knowledgebase
    - *Inst. & DiSSCo Service policies; self assessment classifications (via metadata schema), scores and vis. of policies/services.*
    - *WP5 representation in 7.3 addressed (Julia Prim Reis, MfN)*
- Highlighted wider DiSSCo needs for self-assessment tools
    - *Institutional digital capabilities (WP3)*
    - *Institutional specialisations (D 8.2)*
    - *May be other dimensions to consider too*
- Discussed whether to craft the milestone as a system agnostic design document or tightly integrated the DSpace knowledgebase.
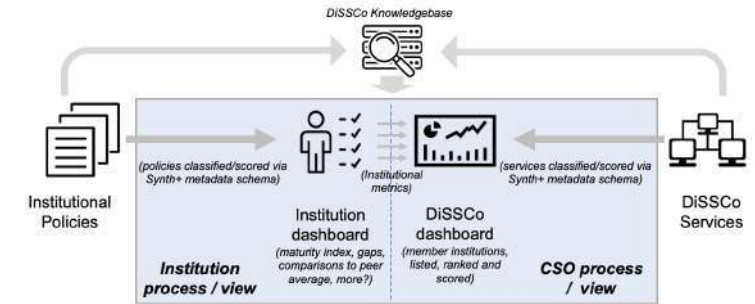
Online checklist tool: user journeys

User stories to design requirements

3

DiSSCo All Hands - Task 7.3

DiSSCo Policy Tool - Mapping and Gaps

Wed 20th January 2021, 13.00-14.30 CET / 12.00 - 13.30 GMT

Notetaker: Josh Humpheries, NHM London

Meeting Chair and Editor: Vince Smith, NHM London

Participants

Vince Smith, NHM London, Task 7.3 lead
Matt Woodburn, NHM London
Patricia Mergen, Meise Botanic Garden
Judite Alves (Museu Nacional de História Natural e da Ciência, ULisboa)
Gergely Babocsay, HNHM, Budapest
Eva Häffner (BGBM)
Tina Loo (Naturalis)
Jiri Frank (National museum, Prague)
Elsa Fontainha (ULISBOA WP 1 T 1.4., member)
Elspeth Haston (Royal Botanic Garden Edinburgh)
Serge Scory (RBINS)
Marie-Laure Kamatali (CETAF)
Quentin Groom (Meise Botanic Garden)
Carole Paleco (RBINS)
Patrick Semal (RBINS)
Jonathan Blettery (MNHN, Paris)
Laura Tilley (CETAF)
Heli Fitzgerald (Finnish Museum of Natural History, Luomus, University of Helsinki)
Marko Hyvärinen (Finnish Museum of Natural History, Luomus, University of Helsinki)
Pedro Arsénio (ULisboa)
Katharina Wölfel (NHM Vienna)

Purpose of session

This session aims to review current progress, outline a vision for the deliverable and develop high-level contents for the sections of the associated milestone (MS7.5 Design of the DiSSCo policy framework tool). This milestone will serve as a design blueprint for the task deliverable. This deliverable is an online checklist tool which allows a DiSSCo Partner to map their institutional policies against the policy requirements of DiSSCo Services to show policy alignment, and for the DiSSCo CSO to see the overall state of policy compliance and gaps across all DiSSCo Partners. The tool will support the upload/linkage of institutional policies; deposition of a list of the policy requirements from the DiSSCo CSO for the DiSSCo

Services; contain a classification of terms (metadata schema) for these policies/services; and finally, the self-assessment interface that allows a user to apply the metadata schema and classification terms to their institutional policies, such that they can demonstrate / self certify alignment with the DiSSCo service policy needs.

The session will draw on a series of user stories previously compiled through 7.3 meetings.

Main discussion, highlights and decisions

(Full original minutes / notes are at the end of this document)

13:00 - 13.05 (CET) Welcome and introductions
(See participant list)

13:05 - 13:15 Presentation of the task overview and progress to date (agenda items 1 & 2)
(See Action 1)

13:15 - 13:30 Presentation and discussion on the vision for the deliverable (3)
(See Action 2 and 3)

13:30 - 14:00 From user stories to definitions, requirements and a design blueprint (4)
       - Review and discussion of the policy tool requirements worksheet
       - Discussion on converting the user stories into a set of requirements
       - Review "next steps" (outstanding tasks) slide
(See Action 4)

14:00 - 14:15 Discussion of the milestone report structure (5)
       - Are these sections right?
       - What is missing?
(See Action 5)

14:15 - 14:25 Revising discussion points (6)
       Does the vision for the policy tool align with needs / expectations?
       - For example, this aligns with the need for self-assessment tools in WP3?
       Are the tasks (next slide) the right ones & are we missing anything?
       Are the section headers for the milestone correct?
       - Are there better blueprint examples we might base this milestone on?
(See Action 6)

14:25 - 14.30 AoB and close

Actions, Next Steps and Conclusions

1. AGREED: We Recognised the complex policy landscape (many policies incomplete for both institutions and DiSSCo Services)
2. AGREED: We confirmed that a self assessment framework was the right approach

3. AGREED: We Discussed alignment with SYNTH+ 2.1 (policy metadata schema) and that the misalignment of timelines can be managed.
4. AGREED: We discussed points of integration with WP5 knowledgebase and CONFIRMED that these related to 1) Institutional and DiSSCo Service policies; 2) self assessment classifications (via metadata schema), scores and visualisations. of policies/services.
5. ACTION: We determined that WP5 representation in 7.3 addressed need to be improved and have addedJulia Prim Reis, MfN to address this need.
6. DISCUSSED: We highlighted a number of instances where there are wider DiSSCo needs for self-assessment tools. In particular these relate to Institutional digital capabilities (WP3), Institutional specialisations (D 8.2), and noted that there may be other dimensions of DPP to consider as well.

Original Meeting Notes

VS introduction
- Slides are available in the main agenda for the AHM
- Single milestone and single deliverable for the task
- The milestone and deliverable deadlines have both been extended
- 7.3 is not about writing institutional policy nor DiSSCo policy - these tasks are out of scope
    - Do need to identify some high level policy needs for DiSSCo though
- The tool created needs to be useful for both institutions and DiSSCo itself
    - We need to find ways to incentivise people to use the tool
- Eva Haffner asked when the DiSSCo policies will be written if they're not being written in this task.
    - VS: ELViS is an example project where concrete policy is required, other requirements aren't as clear. Likely the CSO that needs to develop/oversee this. The tool needs to be flexible enough to accommodate the policy once it's developed.
    - Ana Casino - Synth task 4 has a task to create metadata for policy, could be useful. Some policy will need to be developed internally to DiSSCo but others will have to be developed by the institutions as it needs more detail and isn't DiSSCo's responsibility to develop that.
    - Helen Hardy - there are some other WPs and tasks (like 3.2) which are developing best practices which could be used to build policy off later.
    - Patricia Mergen - Some institutions are bound by national level policy and rules which they can't change directly. Some of these rules go beyond Europe and make the task of creating policy much more complex
    - VS: we have a comprehensive list of policy areas (produced through ICEDIG) which are being further developed in Synthesys. These include external policy commitments as well (nagoya, fair data etc).
- Progress to date:
    - Tool "statement of purpose" developed
    - 29 outline user stories developed & classified

- Initial policy classification / metadata scheme (developed in ICEDIG & being refined in SYNTHESYS+) - https://bit.ly/3o2TBTz
    - Policy categories in the spreadsheet were extracted from an assessment of policies that existed in 6 ICEDIG partners
    - This is a draft piece of work and is being developed in Synth
    - Laura Tilley - in synth 2.1 the main objective is to find the core policy requirements of all DiSSCo services. ELViS is the focus to start with. Currently working through the schema linked above and working out the core policies to create a set of minimum required policies. This will then be translated into a metadata schema
        - VS: the delivery schedule is slightly out of sync with the timeline of this task which could cause some issues but are overcomeable.
- Exemplar design blueprint documents identified (aiding milestone structure)
- A vision statement for the policy tool has been developed (statement of purpose about what the tool will do) - available in the slides
    - VS has created a visual guide which shows the two most common user journeys
    - The DiSSCo knowledgebase is part of the graphic because it would be good if the outputs of this task were stored in the knowledgebase
        - A lot of institutions policies aren't open currently
    - Quentin Groom - noted that policies are often not open because they're not in a fit state to share them rather than because the contents is private. Perhaps overall summaries can be generated instead of releasing the entire document.
        - VS: we're interested in the classification of the policy, not necessarily the policy itself.
        - QG: however, some use cases will require the entire policy to be visible
    - Elsa Fontainha - how is the internal and external communication policy for DiSSCo embedded in the policies?
        - AC: communication policy strategy is being worked on in wp8. This strategy is not directly related to the policy although it is of course related.
        - VS: clarifies that this tool being developed in 7.3 is about policy covering DiSSCo services.
    - Mareike Petersen - a closed space in the knowledgebase is fine and other use cases being discussed in this task should be achievable.
        - MW: the metadata schema needs to be able to cope with various ways that policy is presented - could be a document, several documents, a url, a person etc
    - Chat comment from QG: *"It takes time to become compliant with policies, but by making a policy open you are expected to conform immediately. This is a good incentive, but is more pressure than many institutions will want "*
    - VS: does our tool need a way of verifying the categorisation through external validation? This needs to be thought on more.

- Question in the chat from Patrick Semal: *"How to manage the multilingal aspects of the policies. How to check if the policy in one language is equivalent to another one?"*
    - In ICEDIG there was a multilingual team who read through the policies. There were still pieces of documentation that couldn't be read and some of these were passed through translation tools. This was fraught with danger but was useful for a simple overview of what the document was about. We need to work on this more for when we scale up to doing DiSSCo level policy parsing.
    - VS: Because the institutions will be self assessing in the first instance, they should be able to read their own language. External validation could be challenging though. The metadata should help make this as easy as possible to represent the policy documents across all languages. VS also notes that the language issue also applies across all the DiSSCo services too (e.g. ELViS).

MW - draft statement of purpose and tool requirements

- Available here: https://docs.google.com/spreadsheets/d/1DOEY35KH0o5MGFaPBbtRtnbB9xTF1CnzwRcqkiqBC6I (draft, work in progress)
- Mix of methodologies included (agile, ad hoc etc)
- A collection of high level user stories (epics) have been collected
    - These have been categorised into a few high level categories
- From the user stories we can then extract:
    - Glossary of terms
    - Roles that relate to the users of the tool (this exercise also shows gaps where expected roles are missing from the user stories)
    - Functional requirements with links back to the business requirements they came from
    - Concrete features of the tool itself
    - Data elements for the backend of the tool (some of these have already been identified in ICEDIG and are being worked on in SYNTHESYS too) which creates linkage to metadata schema
    - User access requirements - the system will be quite open (everyone can see most things, but with restrictions on who can upload documents for which institutions)
    - Reporting requirements - what kind of visualisations will we want to be able to generate?
    - Non-functional requirements - performance, security etc. Need to consider how this tool will live amongst the rest of the DiSSCo architecture.
- VS: We can produce a functional pilot within DiSSCo prepare, but we need to consider what we want to achieve beyond that point.
- VS: Are there others outside of this task group who should be contributing to this document? E.g. DiSSCo CSO/knowledgebase.
    - Mareike Petersen comment from chat: *"We should defiantly bridge a potential gap between the knowledgebase development and the Policy Framework,*

either through Matt or Josh or I could ask if somebody of my Team could join as well"

- **[Action]** VS suggested MP suggest someone to join the WP7.3 group to cover this off.
- AC in the chat: *"@Mareike, I think it will be important that Laura Tilley, of CETAF, joins the team as well, at least to certain extent (she is handling the metadata schema in SYNTHESYS+ for this framework tool). Thanks "*
- MP in the chat: *"@ Ana: If I understood Vince correctly he would like that somebody from 5.1 joins the Task 7.3 Meetings. Maybe Laura should join here as well? At least when it is related to the metadata Schema? "*
- HH in the chat: *"Would be helpful if I or someone from 3.1 could join at least occasionally too "*
- AC in the chat: *"Yes, either under T7.3 or under T5.1 (and towards both at the very end) we should be able to coherently align efforts with the expected outcomes from SYNTHESYS+. Thanks "*
- LT in the chat: *"Yes I usually attend the 7.3 meetings, but can attend the 5.1 also. "*
- Elsa Fontainha: commented that it's a rich source of information, and could be used to forecast users.
    - VS also noted that ELViS will also house a lot of very useful user information. Between that and this we have a rich source of information.
- HH: In task 3.1 meeting there is an appetite for an institutional self assessment tool against a range of digital maturity/transformation. It would be useful to link that to the tool being created in this task because policy is a strong part of that and often an indicator of digital maturity in of itself. This task also has a prototype deadline in April as well.
    - VS: Could creating more self assessment tools be a part of future post-SYNTHESYS projects?
    - HH: We need both high level tools that are almost just a series of questions to consider, which then link through to more in depth tools to do a self-assessment. We can develop more in depth tools after the high level tools.
    - VS: We're at the stage where we are developing prototypes and then we'll develop more feature rich tools further down the line.
    - Carole Paleco noted that the work being done in 7.3 will feed into WP8 too.

VS - from user stories to a design blueprint (the milestone)

- Issues with defining the key concepts and user roles
- Need to continue to convert more user stories into functional requirements
- Need to think about user access requirements
- Need to think about reporting requirements (even DiSSCo CSO haven't thought about this yet in any great depth - confirmed by AC. AC: need to wait a little bit for the metadata, knowledgebase etc links to mature).
    - VS: The pilot will start the process and then we'll work in an agile fashion so that we can respond to the requirements that develop over the course of the project.

- VS notes that this task wasn't originally envisaged as a technical task so there needs to be a bit of management of resources to ensure we have all the resources we need to complete the deliverable.
- We've started basic structuring of the milestone document

VS - revisiting discussion points

- Does the vision policy tool align with needs/expectations?
- Are the tasks the right ones?

MW question: we could create a c (DSpace etc). The latter would get us closer to the deliverable but which would we prefer to do for the April milestone?

- VS: we'll probably end up doing both.

QG question in the chat: *"We have not discussed the machine readability of the data. This would be important to build the policy decisions into services".*
- VS: We need to consider this more and it's not currently covered in the spec docs. Should be covered by the metadata though so we'd just have to expose this through APIs.
- MW: Might feed into a knowledgebase API requirement - other DiSSCo systems might want to visualise data from the knowledgebase.
- AC: These visual tools might need to be joined to a higher level gateway rather than a specific service
- VS: We need to make sure we don't create a large unsustainable set of technical requirements and deployments.

Patricia Mergen question: what are the plans for ELViS post SYNTHESYS? Who will take it over?
- VS: As we move from DiSSCo PREPARE to DiSSCo CONSTRUCT ELViS will move over. The funds should cover that. ELViS may have to function beyond the initially supported date anyway because of the extensions that are occurring due to COVID.
- AC: ELViS has to be transferred to the DiSSCo platform too. Resources will need to be allocated and then secured in DiSSCo once we have moved to DiSSCo being a legal entity.
- VS: there are substantial amounts of funding coming into DiSSCo which you would hope will bridge the gap between DiSSCo PREPARE/SYNTHESYS and other EC projects. We need to decide which projects are picked up and moved over and which are left as pilots that finish with the project they were created within.

Laura Tilley question: the slight misalignment with the SYNTHESYS task 2.1 work - does VS have any suggestions on this?
- VS: move on the 7.3 team to be able to accommodate any metadata schema at any point. The standard should be agreed and then the content can come later.
- MW: Creating the data model in 7.3 and then comparing with the SYNTHESYS task will create a bidirectional flow where we communicate and influence each other.

- VS: There isn't much else we can do because 7.3 can't slow down and 2.1 can't speed up. So we just need to have a good dialog and align our standards as best we can.

ALL HAND MEETING - AHM1
WP8 Wrap-up

Ana Casino - CETAF

DiSSCo Prepare

WP8 Leader

AH1 / Business stream

WP8 Stakeholders engagement and comm strategy
Thursday, 21 January (15h00-16h30) – Ana Casino (CETAF)

Update/discussion on engagement and strategy
on Communication material, advocacy strategy and updates (FF) and NNs engagement

Specific focus on Thematic Specialisation Plan (D8.2, July'22)

- Shared vision on the basic statement (self-assessment tool)
- Definition of domains, areas of influence to be covered
- Identification of a workplan to collate information, analyse and classify it, produce the tool
- Contributions from NNs deemed to be instrumental to acknowledge distributed specialized capacities across partners

Upcoming deadlines: MS8.3 "Collection of institutional strategies and policies" (Feb'21)
MS8.5 "Initial findings for the Specialisation Plan" (Jul'21)

Challenges to address:
- Taylor-made solutions from DiSSCo overall vision useful for national priorities/requirements
- For the Specialisation Plan:
    * Overcomplexity
    * Long-term approach
    * Parallel/complementary web-based endeavor (with e.g. T7.3 Policies framework)

DPP AHM1 WP8 Stakeholders engagement and Communication session

Date: Thursday January 21st, 2021

Time: 15.00 to 16.30 CET

Organizers: Ana Casino and Marie-Laure Kamatali

Presenters:

- Ana Casino and Marie-Laure Kamatali (CETAF)
- Serge Scory and Carole Paleco (RBINS)
- Eva Alonso (Naturalis)

Reference Documents:

- T8.1 session document:
  https://docs.google.com/document/d/1S1EGXqSmYmqIsPtieP3_H8Xqa1_bLATkqul
  P_BbQifo/edit#

Agenda

Time	Topic	Lead
15.00-15.05	Welcome	AC
15.05-15.20	ENGAGEMENT 1.Task 8.1: National Nodes engagement - Description of the task and its context within DiSSCo RI - Objectives of the task - Achievements - Challenges - Timeline - Connected Milestones and Deliverables	AC/MLK
15.20-15.35	2.Strategic engagement - Existing mechanisms - New or improved ones	MLK
15.35-16.05	3.Specialisation Plan - MS8.3 Institutional strategies & policies collected - Presentation of the Specialisation Plan - Open discussion on the uses of collected data	MLK/RBINS
16.05-16.20	ADVOCACY 4. Task 8.4: Advocacy Strategy Implementation of advocacy plan- EA - Funders Forum - Advocacy strategy 2nd stage	EA

16.20-16.25	5. Next steps	AC
16.25-16.30	AOB	AC

Session Type: working meeting.

Notes Taker: Celine Cassarino and Laura Tilley (CETAF).

Participants

- Alex Hardisty (CU) (partial attendance)
- Paul Braun (MnhnL Luxembourg)
- Vince Smith (NHM London)
- Celia Santos (MNCN-CSIC)
- Isabel Rey (MNCN-CSIC)
- Steffen Kiel (NRM, Stockholm)
- Luca Bartolozzi (NHM, UNIFI)
- Niels Raes (Naturalis Biodiversity Center/NLBIF, Netherlands)
- Patricia Mergen (Meise Botanic Garden)
- Eva Häffner (BGBM)
- Tina Loo (Naturalis)
- Jose Alonso (Naturalis)
- Sarah Rossi de Gasperis (NHM, University of Florence)
- Martin Vipp (University of Tartu)
- Heli Fitzgerald (Finnish Museum of Natural History, Luomus)
- Aino Juslen (Finnish Museum of Natural History, Luomus)
- Karol Marhold (IBSAS, Bratislava, Slovakia)
- Piotr Tykarski (UW)
- Rosarosa Manca (NHM, University of Florence)
- Patrick Semal (RBINS)
- Dimitris Koureas (DiSSCo-CSO)
- Eva Alonso (DiSSCo-CSO)
- Wouter Addink (Naturalis)
- Serge Scrory (RBINS)
- Vince Smith (NHM)
- Aino Juslen (Luomus)
- Anne Koivunen
- Boyko Georgiev
- Elsa Fontainha (Lisboa)
- Francois Dusoulier (MNHN)
- Gergely Babocsay (HNHM, Budapest)
- Gianna Innocenti (NHM, University of Florence)
- Heimo Rainer (NHMW)
- Henrik Enghoff (UCPH)
- Hugo de Boer (University of Oslo)

- Jana Hoffman (Museum für Naturkunde, Berlin)
- Jiri Frank (National Museum, Prague)
- Luca Bellucci (NHM, University of Florence)
- Maria João Fonseca (U.Porto)
- Michel Guiraud (MNHN)
- Peter Warth SMNS
- Salomé Landel (MNHN)
- Karol Marhold (IBSAS, Bratislava, Slovakia)

Notes Takers: Céline Cassarino and Laura Tilley (CETAF).

Aim of the session:

Planned agenda:

15.00-15.05	Welcome	AC
15.05-15.20	ENGAGEMENT 1.Task 8.1: National Nodes engagement ● Description of the task and its context within DiSSCo RI ● Objectives of the task ● Achievements ● Challenges ● Timeline ● Connected Milestones and Deliverables	AC/ MLK
15.20-15.35	2.Strategic engagement ● Existing mechanisms ● New or improved ones	MLK
15.35-16.05	3.Specialisation Plan ● MS8.3 Institutional strategies & policies collected ● Presentation of the Specialisation Plan ● Open discussion on the uses of collected data	MLK/ RBINS
16.05-16.20	ADVOCACY 4. Task 8.4: Advocacy Strategy Implementation of advocacy plan- EA ● Funders Forum ● Advocacy strategy 2nd stage	EA
16.20-16.25	Next steps	AC
16.25-16.30	AOB	AC

Recommended material: none

NOTE

ITEM 1 WELCOME by *Ana Casino_CETAF (AC)*

Presentation of the contents of the meeting addressing several issues of engagement and advocacy:

- Strategic engagement activities
- Present the definition and basis of the specialisation plan.
- How we reach our major stakeholders (e.g. the Funders Forum)
- How we are going to address our future National Nodes meeting.

ITEM 2 ENGAGEMENT by AC and MLK

- **Current updates of T8.1:** challenges faced and how to overcome them. It is important for National Nodes (NNs) to have well established contact points at a national level and to channel information both ways (from DiSSCo to National representatives and vice versa), this should be strengthened and well covered.
- **High engagement from the NNs**: they have worked hard to contribute to surveys (country factsheets). NNs are instrumental to the engagement of National Governments so they are plenty aware of the benefits of joining DiSSCo and what it will bring to the respective national environments.
  - we have been able to build a corpus of policies.
  - we have opened a channel of communication trying to provide you with material to facilitate NNs engagement work with governments, institutions and others within your NNs. There are different levels of complexity for each NNs (number of entities composing the NN, the framework & circumstances influencing each node), thus the communication material is instrumental in disseminating the goals of DiSSCo in a more harmonised way.

- **Channels for engagement:**
  - NNs with national governments (vertically)
  - Across the NNs representatives (horizontally) : interactions among the NNs are essential for the specialization plan- it will help build the strategic mapping- so we can better oversee how DiSSCo can be implemented as a distributed research infrastructure.

ITEM 2 ENGAGEMENT by MLK_CETAF

WP8 achievements so far which included:

- Data collection:
  - national priorities transformed into a report and a matrix to identify commonalities and differences in the landscape.
  - the Funders Forum surveys (bilateral meetings) that inform advocacy actions of the NNs.

- Participated in 9 monthly meetings. Very insightful to circulate the updates of DiSSCo and CSO actions.
- Launched an internal engagement campaign (*DiSSCo Happy Hour*) addressed to Communication departments across the institutions.
- T8.2 outcomes: External communication tools developed will be shared at the next NN meeting (Feb 21) , Including DiSSCo Key messages webpage, DiSSCo brochure, DiSSCo ppt.

Challenges & lessons learned:

- During the launch of the internal engagement campaign, we noted that within the communication staff is not engaged at institutional level. So it could be better to try engagement with institutional staff and then scale up the process of engagement.
- During NNs meetings we need to revise the structure so that there is a two way communication rather than just one way (dissco cso communicating to the NNs). Moving forward, suggestion to have a 30min slot for WPs input from NNs (to share challenges they face, success stories,..) but also for NNs consultations about other WPs.
- There needs to be strengthening on clarity of the DiSSCo vision, by better understanding the landscape - we need to think how to develop tools to facilitate clarification. Jose Alonso is the new DiSSCo communications officer and is currently involved in restructuring the DiSSCo website.
- Lack of clarity on what the DiSSCo universe is, which is connected to developing a clear definition of DISSCO vision , how do we make the key messages clear: among that, restructuring the website and co.

Timeline with the Milestones and deliverables for T8.1 (see meeting slides) both the outcomes completed and the upcoming ones.

Way forward:
- Now is the time to incorporate new methodologies for engagement, it is planned to implement a monthly newsletter.
- 30 mins time slots will be incorporated in the next NNs meetings to facilitate two-way communication.

Questions from meeting participants about the mechanism of engagement:

Vince Smith-NHM: communication is usually around a lot of internal matters and if we are trying to reach out to governments who have larger biodiversity interests, then we need to make sure we feed into that, rather than just our own internal interests. Thus we may need to tailor the language and emphasis and tones to different stakeholders. Need for a change in tone and emphasis (focus on "how DiSSCo responds to governments' priorities").

- *AC* added that it is important to describe DiSSCo in the global landscape, and the communication tools should be complementary to each other (e.g. the DiSSCo key messages and the National Priority areas).

Lucas B - agrees that the newsletter could be a good instrument for communication but it should also be addressed to those outside the DiSSCo family.

- *AC* mentioned that CETAF is trying to connect with communication departments in institutions to broaden the outreach. It is a process that CETAF has just started.

Eva Alonso-Naturalis added that it is important that the communication is in alignment with NNs priorities, thus DiSSCo should think carefully in terms of narrative. The funders forum will facilitate our work.

Jose Alonso - Naturalis agreed with Lucas and Vince and admits that DiSSCo needs to increase its role in the global landscape, he feels that has been missing slightly.

- Shared a reference article: [How natural history museums should play a bigger role in finding the sources of wildlife pathogens](#)
- Jose is currently restructuring the DiSSCo website to better provide double pathways of communication, i.e. for those within the community, and more simplified for those outside DiSSCo.

Niels Raes -Naturalis mentioned the importance of engagement with small institutions. For the Netherlands some small institutions have difficulty algning with DiSSCo. He has set up meetings with smaller institutions to inform them of some of the changes taking place.

- *AC:* the experience of the Netherlands may help others with the same complexities, small institutions must have a say in the alignment and acknowledge aims.

ITEM 3 THEMATIC SPECIALISATION PLAN by Serge Scory - RBINS

The specialisation plan is due in 2022. Serge commented that a specialisation plan mentioned in several sections is the DPP grant agreements, in general terms, but an actual definition is missing. There have been various meetings with CETAF to gain a definition.

RBINS has drafted a definition and shared this in the meeting: '*The organizations contributing to DiSSCo form a very rich but diverse network. Consequently, many assets are unevenly distributed amongst these organizations. In order to document, promote and optimize the assets, and possibly to identify gaps, we will develop a tool to collect all relevant information and enable the assessment of the institution's specificities.'*

As a starting point RBINS suggested to **collect information** from institutions:

- Type and size of collections;
- Digitisation techniques and capacity;
- Training (capacity and domains);
- Strategic goals of the institutions (services, policies and willingness to contribute.

Source of collected info: other WPs, Projects and NNs.

T8.2 leaders suggest that the specialisation plan could be displayed as a visual assessment tool. They need to work further the criteria and requirements for such a tool and the technical feasibility for a building such a tool, can technical expertise be provided from other WPs etc → need to analyse through use cases the potential use of this tool.

Questions and Comments regarding the thematic specialisation presentation:

Patricia Mergen_BGM/RMCA: will you also focus on the capacities and technical infrastructure?

Vince Smith_NHM: it is important to understand what the data collected will be used for, what the specialisation plan will be.

Dimitris Koureas_DiSSCo CSO the definition is tricky and something we gradually need to understand. Original purpose of a specialisation plan was to understand the DiSSCo landscape for instance which partners are best to provide service that will reach the users, this is important for knowing how the different parts of the RI can be developed, and how we help institutions or nodes to deliver those services. Also what priorities of the institutions, or countries can we align with.

AC commented that the specialisation plan refers to the distributed resources that we need to put together through a single entry point. We need to Identify the dimensions and domains that the specialisation covers. A group of countries/institutes may have the same expertise. DiSSCo aims to provide distributed services.

Niels Raes_Naturalis The SYNTHESYS+ and the Dutch dashboard of collections are useful for showing specialisation in collections. Niels also mentions an example of how the Dutch collections Dashboard was very useful for showing specialisation of geological collections is a smaller Netherland institution. Smaller institutions also have unique specialisations

Vince Smith_NHM teasing out so many dimensions on what specialisation means. We need a plan around a specialisation rather than more dashboards. There has been some work done in ICEDIG around digitisation activities and what makes sense at institutional, and national level, this may provide useful information. A plan of what a specialisation plan might be more useful and sustainable.

Dimitris Koureas_DiSSCo CSO - A specialisation plan is about understanding the DiSSCo landscape. Long Term perspective why - but we need to understand why we capture the information.

Serge Scory_RBINS: said that we will start the work using the definition as a basis. Other work packages will be screened for expertise and information, and an inventory for partners who have high PMs that can provide user case analysis.

AC one important task will focus on the use cases, this effort might take years, and this is just the starting point, partners may have additional specialisations not covered yet.

Michel Guiraud_MNHM: Agrees that it may take years to develop a specialisation plan. DiSSCO is about offering services, we need to think about what are the questions DiSSCo needs to solve when making a specialisation plan. What are the scientific questions that DiSSCo wants to answer. We only really knowknow the taxonomic questions, we do not know what is really required yet.

Patricia Mergen_BGM/RMCA: Lifewatch have a thematic hub of services, they are an ERIC and operational. They still now have some countries considering to be a thematic hub as new competencies arise. Thus it is good to remain flexible as new specialisations arise.

Dimitris Koureas_DiSSCo CSO mentioned two points: 1) There is always a simpler way of looking into DiSSCo. DiSSCo is an amplifier for the scientific objectives that our collections have. Dissco goes through our own missions, and scales them up. The RI relies on scientific collections, thus we can start from the missions of individual institutions.

2) we need to see the benefits of the mechanics of the institute infrastructures. In Naturalis many of their discussions are about how their activities can complement what is going on externally. We are still not sure of the DiSSCo landscape. We can only look in our institutions as first for small parts of the puzzle.

AC: There are several dimensions of a specialisation plan, if we identify the distributed services then we can build on those specialisations, thus partners will not need to all individually open positions for specific expertise because we can share the expertise. We will have a collation of resources and provision of services, so not everyone has to provide the same service. For those that do have a specialisation, we can build on further. Small institutions can also have specialisations.

Michel Guiraud_MNHM Does not fully agree with what is previously said, and says we should remember that in the beginning we built up DiSSCo saying it will give an added value and should solve questions that individuals can not provide.

Vince Smith: Mentioned that those aspects of the specialisation task remind him of the effort on the one world collection, where they looked at taxonomic expertise. There was a crosscutting effort to link staff expertise and institution collections. Maybe there are lessons that could be taken from, or rather what not to do. May help get more understanding. NHM has the data from this but is not the owner.

Patrick Semal_RBINS: It could, but we also have a lot of existing players in this field that we need to take into account.

AC said we will keep sharing this at NN meetings.

ITEM 4 ADVOCACY STRATEGY BY EA

EA explains why it's important to start advocacy actions with the National Funders forum advisory board.

This task includes:
- implementing the **advocacy strategy** including advocacy actions with **an objective: to reach effectively optimal coordination & alignment with national authorities.**

Advocacy key actions go together with other tasks to collect and analyse data from nodes. Therefore, the communications tools that are being developed are essential to support advocacy actions.
- **Liaise effectively** with national authorities to ensure their practical commitment & alignment with national priorities

→ **First Funders Forum gathering will take place in February 2021**

Key Actions
- ▢ Collecting and analysing information provided by NNs (NNs consultations)
- ▢ Advocacy Strategy > Phase I – July 2020
  - ▢ **FF information package** (online/hard copies) & support letters
  - ▢ **Tailor-made national advocacy actions**
    - ▢ Bilateral meetings (>21),
    - ▢ Peer to peer advocacy (NWO – The Netherlands)
    - ▢ Country fact-sheets
- ▢ Supported by communication tools (key messages, social media, brochure, etc.)

Bilateral meetings
- allowed us to have **a better understandin**g of the national priorities - national fact-sheets,
- support **common development of national advocacy actions** - country specific,
- **trigger regional discussions** (Scandinavian countries) // discussions on DiSSCo position

Peer to peer advocacy actions:
- To facilitate national engagement (UK case)
- Led by the Dutch Research Council (NOW)

Why Funders Forum is so important, as an advisory board, it will:

- **inform** on matters related to legal and financial recommendations of DiSSCo RI
- **enable DiSSCo to effectively adjust** its development to national and international priorities
- **facilitate the key stakeholders endorsement** of the core implementation and operational principles
- **enable the national representatives to make informed decisions** on their future financial commitment to DiSSCo and its relationship with the scientific institutions
- enable a more consolidated degree of **engagement and cooperation between DiSSCo** and the scientific institutions

Key achievements:

- ☐ DiSSCo awareness in 21 countries
- ☐ Funders Forum Advisory Body inaugural meeting
    - ☐ Online meeting on February 25th.
    - ☐ Current representation (8 countries) - BELGIUM, GREECE, U.K, SLOVAKIA, DENMARK, BULGARIA, ESTONIA, THE NETHERLANDS
    - ☐ +3: FRANCE, ITALY, LUXEMBOURG
    - ☐ In discussions: PORTUGAL, FINLAND, SPAIN
- ☐ Together we are building an even stronger community based on a common understanding

Next steps:

- NNs monthly meetings
    - ○ Feb'21: Presentation of Communication material
    - ○ March'21: Review of FF meeting
- WP outcomes:
    - ○ MS8.3-Strategies collected (Feb'21),
    - ○ MS8.5-Specialisation Plan initial findings (Jul'21)
- Launch of T8.3 Stakeholders Engagement
- Towards iGA3 and WGs (Strategic positioning and Concept)
- Close linkage with WP7 (LE and Governance)

ITEM 5 Next steps by AC

NNs monthly meetings focus

- February:
    - ○ meeting will focus on DiSSCo external communication tools.
    - ○ Milestone MS8.2 institutional strategies and policies collected, report
- Milestone MS8.3 Findings from specialisation plan initial findings
- Launch of T8.3 Stakeholders engagement thanks to new people hired at NHM.

- Launch work towards the general assembly in june in italy IGA.
- Close collaboration with WP7.

Zoom Chat

15:01:55 From Laura Tilley to Everyone : https://docs.google.com/document/d/1CFBMUsc0q-F-19wLEamr-GCHW4m27p8vc0h6UUSa0Oo/edit

15:02:17 From Marie-Laure K. to Everyone : https://docs.google.com/document/d/1l-OvOLouGCecUqmBw06OVT56gl5yX5Jqc_CB7saAMV0/edit#heading=h.iy94aivw8mk0

15:02:20 From Laura Tilley to Everyone : Document above please state attendance

15:32:51 From jose.alonso to Everyone : https://www.washingtonpost.com/science/virus-animals-to-humans-specimen-preservation/2021/01/14/2b3c0472-55e9-11eb-a931-5b162d0d033d_story.html

15:39:44 From Steffen (Stockholm) to Everyone : is there an update on the Digital Transformation document that was sent around in October/November last year?

15:40:40 From Elsa Fontainha (ULisboa) to Marie-Laure K.(Direct Message) : @Steffen request.

15:44:17 From Niels Raes to Everyone : The collections dashboard (Synthesys+) will provide information on the type and size of collections

15:45:47 From Elsa Fontainha (ULisboa) to Everyone : @Niels, thanks.

15:46:29 From Wouter Addink to Everyone : Hi Steffen, the received feedback on the digital transformation document has been discussed and will be taken into account in the next steps. These next steps will be worked on in the coming weeks.

15:46:58 From Niels Raes to Everyone : WP1 has collected a large number of user stories, which is a good starting point

15:47:50 From Dimitris Koureas to Everyone : hand

15:51:47 From Elspeth Haston to Everyone : It might be interesting to take one example and work it all the way through as a demonstration and clarification for the wider community for whom it's not so clear?

15:53:06 From Steffen (Stockholm) to Everyone : thanks Wouter!

15:53:24 From patrick Semal to Everyone : Is this specialisation plan targets also the developement of the Collections Management System(s) ?

15:54:52 From Dimitris Koureas to Everyone : @Patrick: It could, but we also have a lot of existing players in this field that we need to take into account.

15:54:53 From Wouter Addink to Everyone : identifying the unique value of a collection or organisation may help in identifying what it should specialize in.

16:10:11 From Dimitris Koureas to Everyone : I agree with that Michel

16:22:40 From Karol Marhold to Everyone : I guess Slovakia should be there

16:23:07 From Dimitris Koureas to Everyone : Our apologies Karol !!!

16:23:28 From   Francois DUSOULIER (MNHN)   to   Everyone : France is engaged and will participate. Only the names of the Ministry représentatives are still unknown

16:29:00 From Dimitris Koureas to Everyone : Thank you Luca !

16:31:42 From Jiri Frank to Everyone : Thank you for a great session.

16:32:07 From Vince Smith, NHM to Everyone : Apologies - I have to go - Bye all

16:32:19 From Jiri Frank to Everyone : Bye Vince

16:32:20 From Jana Hoffmann to Everyone : I received it. Thank you!

16:32:21 From Carole Paleco to Everyone : bye thanks !

16:32:29   From   Dimitris   Koureas   to   Everyone   : https://us02web.zoom.us/j/81927609291?pwd=WEpSa0l6VWtqbExYODBSSHJnUWNoQT09

16:32:42 From Jiri Frank to Everyone : Done thanks :D

16:32:43 From Patricia Mergen to Everyone : Got it fine thanks

Agreements / outcomes:
- Session agreed task meeting frequency every 4 weeks
- Subtasks discussed and agreed by Task Team
- Best practice procurement work will take place in Task 4.4

Actions / next steps:
- Meeting and discuss overlap with SYNTH+ WP5/NA5
- Review and integrate strategic positioning from General Assembly work
- Re-write subtasks 8.3.2 and 8.3.3 – split stakeholder analysis and 'road test'

Open Issues:
- Balancing specialised financial/accounting work
- Ensure there is a strong focus on consumers/service users including policy makers
- Shift balance from academic to marketing/commercial
- Consider thematic approaches to marketing and stakeholders

Task 8.3 AHM Meeting Agenda & Notes

Background

See linked discussion document

Participants

Laurence Livermore (NHMUK)
Helen Hardy (NHMUK)
Vincent Smith (NHMUK)
Patricia Mergen (MeiseBG)
Salomé Landel (MNHN)
Mathias Dillen (MeiseBG)
François Dusoulier (MNHN)
Piotr Tykarski (UW)
Tina Loo (Naturalis)
Sarah Rossi de Gasperis (NHM, University of Florence)
Lorenzo Cecchi (NHM, University of Florence)
Aino Juslén (Luomus, University of Helsinki)
Marie-Laure Kamatali (CETAF)
Ana Casino (CETAF)
Eva Häffner (BGBM)
Pedro Arsénio (ULisboa)
Carole Paleco (RBINS)
Rosarosa Manca (NHM, University of Florence)
Jiri Frank (National museum, Prague)
Frederik Berger (MfN)
Luca Bartolozzi (NHM UniFI)
Eva Alonso (Naturalis)
Anne Koivunen (Luomus)

Minutes/Notes

LL gave introductory slides [link], including revised timeline with milestones slightly pushed back, links to other tasks/WPs, subtasks and resources.

VS gave an overview of DiSSCo General Assembly task force work on strategic positioning (led by Aino Juslen - Luomus). A questionnaire for infrastructure providers to fill in looking at their coverage (scope, categories) and maturity/phases, to visualise how infrastructures interact (work in progress). This includes GBIF, CETAF, IBOL, iNaturalist, GeoCase,

DiSSCo itself, BHL, CoL and a notional collections-based institution. Very complex landscape with many types of interactions. More work underway on scoring. One aspect to inform this task. Intended to be complete before the next DiSSCo GA in June.

LL Shared the discussion document (linked at the top of this document). **First milestone/subtask 8.3.1** due April 2021 which will focus on procurement strategy and policy - not to level of operational detail but high level policy and principles with examples.
Will need a more detailed understanding of the likely evolution of DiSSCo needs over time.
May need some positioning statements for key areas such as how to involve smaller, innovative suppliers (other examples in doc).
Also needs to take into account both central DiSSCo and the national nodes.
There is a dependency on the Legal entity model where timing may be an issue.
Collaboration with other Research Infrastructures is likely to be useful.

DK - Care is needed about the balance between financial/accounting elements (specialised) and the wider understanding of lines of development /work that we may prefer to progress in-house vs where we may look to external suppliers. This latter part is also strategic and very relevant at this stage, leaving some of the economic details for later input when the legal entity is understood, as these determine the procurement framework overall.

PM noted that aspects may be relevant to 4.4 - the discussion with other ESFRIs in particular (via the consultant during 2021). PM also is in touch with US. 4.4 Milestone is due in July 2022 and final deliverable is due Jan 2023.

EA - do the milestone changes have knock-on effects to other components?
LL - may have to revisit some aspects when key decisions are made
AC - could limit the scope of the April milestone to allow for later input from the ESFRI interviews etc **AGREED**

LL - the **next subtask/milestone component (now due Nov 2021) is the stakeholder analysis** - not about running fora but agreeing how to engage; who is in scope etc - an analysis of stakeholders maybe with a RACI. Requires the timeline from the first task. Will draw on communication and dissemination strategy. Also includes consideration of the DiSSCo GA taskforce infrastructure services analysis, which will be done in time to inform this.

Will this include user stories? At a high level. It would be good to get user stories verified by users and to develop user groups and get user feedback on nascent services in future.

VS - a lot of work on user stories and users, how does this change what we do?
LL - what about user acceptance testing for services? Ensuring good user experience and interfaces.
VS - we are as much users and consumers as providers of these services?
Good to look at bigger / biodiversity policy users of our data e.g. IPBES, EU policy areas etc? This may be more gamechanging?
May be outside the resources of this task? But may be worth prioritising this group of stakeholders given previous work on other aspects, at a high level.

DK - this was also mentioned at the last GA, and why IPBES were the keynote topic for this All Hands meeting. Work in Synthesys+ including the recent paper are relevant (https://riojournal.com/article/62361/). Starting conversations with Director of biodiversity knowledge centre and what the requirements are for those bodies and data collection - how should DiSSCO present data to be available to those bodies? Not straightforward - great ignorance of the existence of research infrastructures and how they can link information. Complex activity - perhaps this task could consider a particular stakeholder? Otherwise likely too much for scope of this task.

AC - two levels of engagement. Both those who make decisions about using our data at policy/high level, and those with whom we may have practical partnerships not just dialogue - the latter may start from a higher threshold of awareness / be easier to build bridges to?

LL - are relevant organisations in the infrastructure analysis work (of the latter type mentioned by AC)?

DK - yes, up to a point. Perhaps less around making data digestible to policy pipelines? Nothing at present so no tangible case studies of what RIs can provide.

VS - policy users use working groups to perform meta-analyses of publications rather than directly using data? Nothing to plug a pipeline of data into - though worthwhile exploring how that could work in future. These are not the groups that underpin the business case for DiSSCo, which is a bigger efficiency/access gain?

LL - business case also between the national nodes and their governments? Another axis of involving policy?

AC - project [bicycle ?] is a relevant example of how partnerships can be built

QG - this is really a marketing exercise for DiSSCO. We talk to ourselves a lot but others are unaware and we need a marketing mindset to this. ELIXIR for instance have reached out to us around their own biodiversity theme - should we have themes going out to relevant stakeholders? E.g. those who need vouchering of genetic material?

LL - The **final sub task 8.3.3. concerns how we we work with stakeholders**. May include fora, workshops, testing, product fairs, business or academic or policy events etc. What will the channels and methods be?

May need a stronger marketing element?

VS the Synthesys paper (https://riojournal.com/article/62361/) is a possible model to validate findings from a desk based analysis? Could pick a subset of organisations to discuss the findings of an initial piece - cuts across this and the previous subtask.

QG - for IPBES, need to get to the lead authors. Klaus within DiSSCo is connected to IPBES. Could be a good use of short time allocations to task partners.

HH - sub tasks do overlap - have a traditional engagement plan that analyses stakeholders, channels and aims/key messages but then 'road test' this with some exemplar stakeholders/groups? That could also be a timeline stages for milestones i.e. to do a desk-based stakeholder engagement plan and then follow it with real testing/early engagement with key groups?

LL - this seems to fit with QG comments and overall timeline. **Action - discuss at wrap up meeting for All Hands and take forward accordingly.**

Next steps:
- Set task team meetings - **AGREED** every 4 weeks

- Next 3 months focus on 8.1.3 as outlined above
- Review wider timeline for DiSSCo and implications for this task
- Review what doing in the other milestones/subtasks and plan accordingly.
- Best practices as discussed above likely in WP4 but ensure this is integrated properly

AC - meet with GBIF? Work with many stakeholders who we have discussed, also with the Synthesys work. **Action - LL to follow up this suggestion and include NA5 (synth) in next meeting of this task.**

VS - need to classify stakeholder groups early on? Have an output for this, maybe a metadata schema / at least a classification. A schema would help future analysis.
HH/LL - Yes this analysis key in second subtask.
LL - has there been discussion in DiSSCO about CRM/customer side? Answer - not yet.

AJ - changes to staff in Luomus. Both AJ and AK and a new project coordinator will need to be included in this task for messages. **LL to action**
FD - This has made the task clearer, though not yet sure how to apply small resource allocations to this. Look forward to further discussion.

Annex

1. Agenda
2. List of Participants
3. Presentations

DiSSCo Prepare - All Hands Meeting 1 (January 18-22, 2021) - Agenda

Registration in advance is mandatory. Please register [here](#)

Link to the Stream sessions		**[Business](#)**	**[Science & Technology](#)**	
Date	**WP/Task n.**	**Title of the working session**	**Description**	**Convener**
Day 1 (Monday) 18th January				
Plenary 9:00-10:30		Welcome & State of play		Dimitris Koureas/Eva Alonso
		Keynote	Title: Needs and opportunities for natural science collections in Biodiversity 2030	Prof. Isabel Sousa Pinto (MEP-IPBES)
		Break 15'		
10:45 -12:15		Stream sessions	Parallel sessions	Wouter Addink/Ana Casino
		Lunch 45'		
13:00-14:30	WP3. T3.2	Kick-off meeting	Discuss and agree Task 3.2 workplan and subtask leadership. Review related tasks. https://docs.google.com/document/d/12ywJGyo2_4ps8ogBwJ0HP-arVIWId-BpZ8tmZMd1LvA/edit?usp=sharing	Laurence Livermore
		Coffee break 30'		
15:00-16:30	WP7 T7.1	Governance model - requirements and landscape analysis		Eva Alonso/Dimitris Koureas
Day 2 (Tuesday 19th January				
9:00-10:30	WP1 T1.1 & T1.2	Use cases and User stories	Preparation of D1.1 and D1.2, due M 15	Aino Juslén/Mareike Petersen
		Break 15'		
10:45 -12:15	WP1 T1.3 & T1.4	Criteria for prioritisation of digitisation and indicators of socioeconomic impact	Start-up discussion for these tasks which start in M 13	Henrik Enghoff/Rui Figueira
		Lunch 45'		
13:00-14:30	WP3. T3.1 & T3.3	Competencies and capabilities - emerging thoughts	Key conclusions of the analysis (milestone 1); dashboard discussion and 3.3 planning. Support document at: https://docs.google.com/document/d/1V-jPSDg07iweNu-znEHRu1n7f1L6UhukbN_ZvYidy18/edit?usp=sharing	Helen Hardy
		Coffee break 30'		
15:00-16:30	WP5 T5.1	DiSSCo Knowledgebase	Presentation of current version, feedback and requirements by project partners	Sabine v. Mering / Julia Pim Reis / Falko Glöckler / Mareike Petersen
Day 3 (Wednesday 20th January)				

9:00-10:30	WP6 T6.2	open DS I	Recent progress of the development of the Digital Specimen standard (openDS) providing the fundamental data model for the DiSSCO Digital Architecture (DS Arch)	Alex Hardisty/Wouter Addink/Sharif Islam/Claus Weiland
Break 15'				
10:45 -12:15	WP T6.2	open DS II	[continued]	[continued]
Lunch 45'				
13:00-14:30	WP T7.3	Policy Framework (Self assessment) Tool	Final review of high level functional and non functional requirements (from user stories) for the self assesment policy alignment tool, along with an initial design discussion. Supporting document at: https://docs.google.com/document/d/199dKlQrt4NlLLksT5D_SY-J4MIClyczV0__AgtFuTiw/edit?usp=sharing	Vince Smith/Matt Woodburn
Coffee break 30'				
15:00-16:30	WP7 T7.2	DiSSCo Legal Entity	Presentation of and discussion on the Legal Entity chosen	Serge Scory/Carole Paleco
Day 4 (Thursday 21st January)				
9:00-10:30	WP4 T4.1	Methodology assessment	Presentation of MNHN results and discussion about strenghth and weaknesses of methodology	Michel Guiraud/Salomé Landel/Eva Perez/François Dusoulier
Break 15'				
10:45 -12:15	WP5 T5.4	Task planning workshop and GeoCASe/CoL+ roadmap update	Update on GeoCASe and CoL+ roadmaps, and discussion of subtasks and partner roles	Matt Woodburn
Lunch 45'				
13:00-14:30	WP4 T4.2	Scoping exercise for costbook sheet		Laurence Livermore
Coffee break 30'				
15:00-16:30	WP8 T8.1	DiSSCo NNs	Formulate mechanisms for engagement and feedback on implementation of engagement strategy	Ana Casino/Marie Laure Kamatali
Day 5 (Friday 22nd January)				
9:00-10:30	Project Council meeting 1			PC members
Break 15'				
10:45 -12:15	WP2 T2.1	Training strategy	KoM for planning and task distribution	Ana Casino/Marie-Laure Kamatali/Judite Alves
Lunch 45'				

13:00-14:30	WP8 T8.3	Procurement strategy and policy	Working group on MS8.4 in prep for April delivery of the miltestone to set the criteria for the procurement framework	Laurence Livermore/Vince Smith
Coffee break 15'				
14:45-16:00	Wrap-up session	Conclusions		Dimitris Koureas
Time Zone: Amsterdam				

DiSSCo PREPARE AHM1 (18-22 January, 2021) - PARTICIPANT LIST

First Name	Last Name	Affiliation	Role
Agnes	Kirchhoff	Freie Universität Berlin, Botanic Garden and Botanical Museum Berlin	Team member
Aino	Juslén	Luomus	Team member
Alex	Hardisty	Cardiff University	Team member
Alexander	Weigand	MNHNL	Team member
Alexandra	Marçal Correia	Ulisboa	Team member
Alexandra	Cartaxana	Universidade de Lisboa	Team member
Ana	Casino	CETAF	WP Leader
Ann	Bogaerts	Meise Botanic Garden, Curator	Team member
Anne	Koivunen	Luomus	Team member
Anne	Nivart	MNHN	Team member
Anne-Sophie	Archambeau	MNHN/GBIF France WP2	Team member
Anton	Güntsch	Freie Universität Berlin, Botanic Garden and Botanical Museum Berlin	Task Leader
Beata	Papp	HNHM	Team member
Bernadett	Döme	Hungarian Natural History Museum	Observer
Boyko	Georgiev	IBER-BAS (Bulgaria)	Team member
Bruno	Ribeiro	Museu Nacional de História Natural - ULisboa	Observer
Carole	Paleco	RBINS	WP Leader
Celia	Santos	MNCN-CSIC	Team member
Celine	Cassarino	CETAF	Team member
Claus	Weiland	Senckenberg	WP Leader
Dimitris	Koureas	Naturalis	Coordinator

David	Fichtmueller	Botanic Garden and Botanical Museum, Berlin, Germany	Team member
Dulce	Domingos	Universidade de Lisboa	Team member
Elsa	Fontainha	DiSSCo	Team member
Elspeth	Haston	Royal Botanic Garden Edinburgh	Team member
Esko	Piirainen	Luomus	Team member
Eva	Alonso	Naturalis	WP Leader
Eva	Häffner	Botanischer Garten und Botanisches Museum Berlin, Freie Universität Berlin	Team member
Falko	Glöckler	Museum für Naturkunde Berlin	Task Leader
François	Dusolier	MNHN/RECOLNAT/DiSSCo France	Team member
Frederik	Leliaert	Meise Botanic Garden	Team member
Frederik	Berger	Museum für Naturkunde Berlin	Team member
Gergely	Babocsay	Hungarian Natural History Museum	Team member
Gianna	Innocenti	Natural History Museum, Florence University	Team member
Gildas	Illien	MNHN	Team member
heimo	rainer	NHM Vienna	Team member
Helen	Hardy	NHM London	Task Leader
Heli	Fitzgerald	Luomus	Team member
Henry	Engledow	Meise Botanic Garden	Observer
Hugo	de Boer	UiO-NHM	Team member
Inês	Pinto	MUHNAC-ULisboa	Observer
Isabel	Calabuig	University of Copenhagen (Natural History Museum of Denmark)	Team member
Isabel	Rey	MNCN-CSIC	Team member
Jana	Hoffmann	Museum für Naturkunde Berlin	Team member

Jiri	Frank	National museum, Prague	Team member
Jonas	Grieb	Senckenberg SGN	Team member
Jose	Alonso	Naturalis	Team member
Josh	Humphries	Natural History Museum, London	Team member
Julia	Pim Reis	Museum für Naturkunde	Team member
Kari	Lahti	FMNH / UH	Team member
Karol	Marhold	National Node Slovakia - Plant Sci. and Biodiversity Centre SAS	Team member
Katrin	Vohland	Naturhistorisches Museum Wien	Team member
Larry	Lannom	CNRI	Observer
Laura	Tilley	CETAF	Team member
Laurence	Lannom	CNRI	Observer
Lorenzo	Cecchi	NHM Florence (Italy)	Observer
Luca	Bellucci	Museo di Storia Naturale - Università di Firenze	Team member
Luca	Bartolozzi	Museo di Storia Naturale, Università di Firenze	Team member
Maarten	Trekels	Meise Botanic Garden	Team member
Magalie	Castelin	SYNTHESYS+ (MNHN)	Observer
Mareike	Petersen	Museum für Naturkunde Berlin	WP Leader
Maria	Marschler	NHM Vienna	Team member
Maria	Santos	Ulisboa	WP Leader
Maria João	Fonseca	University of Porto - Natural History and Science Museum	Team member
Maria Judite	Alves	ULisboa	WP Leader
Marie-Laure	Kamatali	CETAF	WP Leader
Marko	Hyvärinen	Hyvärinen	Observer

Martin	Vipp	University of Tartu	Team member
Mathias	Dillen	Meise Botanic Garden	Task Leader
Matt	Woodburn	Natural History Museum, London	Task Leader
Michel	Guiraud	MNHN	WP Leader
Mikko	Heikkinen	Luomus	Team member
Mil	de Reus	Naturalis	Team member
Nele	Van der Schueren	DiSSCo Flanders, Botanical Garden Meise	Observer
Niels	Raes	DiSSCo-NL	Team member
Nikolaj	Scharff	Natural History Museum of Denmark	Team member
Ohad	Graber-Soudry	X-officio	Observer
Patricia	Mergen	Meise Botanic Garden	Task Leader
Patrick	Semal	RBINS	Team member
Paul	Braun	MnhnL Lxuembourg	Team member
Pedro Miguel	Ramos Arsénio	Instituto Superior de Agronomia	Team member
Peter	Hollingsworth	Royal Botanic Garden Edinburgh	Team member
Peter	Giere	Museum für Naturkunde Berlin	Team member
Peter	Warth	SMNS Stuttgart	Observer
Philippe	Loret	Museum National d'Histoire Naturelle (MNHN)	Team member
Pierre-Yves	Gagnier	MNHN Paris	Team member
Pieter	Huybrechts	Meise Botanic Garden	Team member
Piotr	Tykarski	University of Warsaw	Team member
Quentin	Groom	Meise Botanic Garden	Team member
Robert	Cubey	Royal Botanic Garden Edinburgh	Team member

Robyn	Drinkwater	Royal Botanic Garden Edinburgh	Team member
Roger	Hyam	Royal Botanic Garden Edinburgh	Team member
Rosarosa	Manca	Earth Sciences Department, University of Florence	Team member
Rui	Figueira	Imstituto Superior de Agronomia, Universidade de Lisboa	Task Leader
Sabine	von Mering	Museum für Naturkunde Berlin	Team member
Salomé	Landel	MNHN	WP Leader
Sarah	Rossi de Gasperis	Museo di Storia Naturale, Università di Firenze (Natural History Museum, Unive	Team member
Serge	Scory	RBINS	WP Leader
Sharif	Islam	DiSSCo CSO	Team member
Sofie	Meeus	Meise Botanic Garden	Observer
Sofie	De Smedt	Meise Botanic Garden	Team member
Steffen	Kiel	NRM Stockholm	Team member
Tania	Walisch	MnhnL	Team member
Tea	von Bonsdorff	Luomus	Observer
Thomas	Winter	BHL	Observer
Tim	Robertson	GBIF	Observer
Tina	Loo	Naturalis	Team member
Vanessa	Demanoff	MNHN	Team member
Veljo	Runnel	University of Tartu	Team member
Ville-Matti	Riihikoski	Luomus	Team member
Vincent	Smith	Natural History Museum, London	WP Leader
Wesley	Tack	Meise Botanic Garden	Observer
Wouter	Addink	Naturalis	Team member

Session:

Task 1.1 and 1.2 "Use cases and user stories"

DPP All Hands Meeting - 19 January 2021

Welcome!

Agenda:

Welcome and plan for the session

Introduction and current status of tasks

Preparation of condensed list of functional demands

Wrap-up and next steps

Convener: Aino Juslén (Luomus) / Mareike Petersen (MfN)

Co-Convener: Sabine von Mering (MfN) / Heli Fitzgerald (Luomus)

Working and Common Notes Document

https://docs.google.com/document/d/1q9wfiGbnKVaj7IFcpt23QbXG-hzu5xEuoLtPQ1hrD1U/edit

Notetaker: Eva Häffner

Zoom Master: Mil de Reus

Introduction and Update on Task 1.1 & 1.2

- Joint MS Report on Use Cases and User Stories
  - submitted and accepted
  - Final Version available via Teamwork: https://dissco.teamwork.com/#/files/9103022

- Use cases follow: „As a [position]… I want to… So that I can …". (adopted from ICEDIG)

1. Research (academic, non-academic incl. Citizen Science)
2. Collection management
3. Technical support (IT & IM)
4. Policy (institutional, national & international)
5. Education (academic & non-academic)
6. Industry
7. External (media & empowerment initiatives)

Introduction and Update on Task 1.1 & 1.2

Use case category	Number of collected use cases/user stories	
	Life sciences	**Earth sciences**
1. Research	271	38
2. Collection management	173	48
3. Technical support	21	4
4. Policy	81	11
5. Education	27	12(10)
6. Industry	19	11 (6)
7. External	5	2 (1)
Total no. of use cases	**597**	**126(118)**

Numbers in () = new Use Cases collected in DPP

Introduction and Update on Task 1.1 & 1.2

Functional demands

- Report to technical WPs
    - What kind of interface is needed?
    - What format?
    - Which type of data?

- Deliverable due by April 2021

In this session...

Goal: Preparation of condensed list of functional demands

- Split into 4 break-out Groups and work on a provided table
    1. Research
    2. Collection Management
    3. Technical support / Policy
    4. Education, Industry and External

- Translate user needs into a more technical requirement (GoogleSpreadsheet)
  https://docs.google.com/spreadsheets/d/1u20k1Exb3KqwCXi_Dx-TVzEQoT_H5Iqi/edit#gid=815645717

- Provide short summary by the end of breakout session (GoogleDoc)
  https://docs.google.com/document/d/1q9wfiGbnKVaj7IFcpt23QbXG-hzu5xEuoLtPQ1hrD1U/edit

DPP AHM1 - Task 1.1 and 1.2 "Use cases and user stories"

Task 1.4: Develop indicators of socioeconomic impact - starts Feb 1st, 2021, due Jan 30, 2023

Deliverable 1.4: Report on socioeconomic impact indicators

Report on socioeconomic impact indicators of DiSSCo and DiSSCo-enabled research and research application

Milestones - MS1.4 Corpus of previous studies on socioeconomic impact compiled WP 1 (M25, Report submitted to EB)

Socio-economic impact

Why perform?

- demand for understanding and evaluation of the return on investment of these facilities to support informed decision-making
- provides RI management with useful information for negotiations with funder

OECD. Reference framework for assessing the scientific and socio-economic impact of research infrastructures. (2019) doi:https://doi.org/10.1787/3ffee43b-en

Socio-economic impact

Why perform?

- one of the dimensions of the Scientific Case assessment of the ESFRI Roadmap

	PHASE			
	DESIGN	PREPARATION*	IMPLEMENTATION**	OPERATION
SOCIO-ECONOMIC IMPACT	– relevance to societal challenges identified and potential economic impact predicted including innovation aspects	- case for impact made, supporting innovation, other types of benefits such as services for society, cultural aspects and attraction of business, industry and public services, etc.	– socio-economic impact cases emerging – capacity building impact proven – contributing to tackling the societal challenges – innovation oriented activities agreed – ability to develop an open innovation culture established	– impact demonstrated consistently – new communities involved – innovation oriented activities operational – private users involved – policies on key societal challenges, e.g. climate change, influenced

ESFRI Roadmap 2021 Public Guide.
https://www.esfri.eu/sites/default/files/ESFRI_Roadmap2021_Public_Guide.pdf

Socio-economic impact

impact ≠ performance

- performance - efficient use of resources
- impact - transformative effect of the RI

OECD. Reference framework for assessing the scientific and socio-economic impact of research infrastructures. (2019) doi:https://doi.org/10.1787/3ffee43b-en

Socio-economic impact

Challenges

- difficult to perform in cutting edge fields
- RI target multiple stakeholders
- research outcomes uncertain and non-linear
- time lag of research
- impacts can be direct and indirect
- change during the lifecycle of the RI
- societal impact may be broad and difficult to capture

OECD. Reference framework for assessing the scientific and socio-economic impact of research infrastructures. (2019) doi:https://doi.org/10.1787/3ffee43b-en

Socio-economic impact

Impact assessment model - reference framework

- socio-economic impact is **broader**: cultural, educational, economic and social
- **users** of the reference framework: RI management and stakeholders
- should be **connected to strategic objectives and mission** of the RI
- useful, with **easy to measure indicators**, user-friendly, reliable and meaningful
- **economic impacts** from commonly recognised indicators
- **social/societal require in-depth validation** of coupling with narratives

Based on surveys and case studies

- Core Impact Indicators (25), organised in 7 strategic objectives and 5 dimensions
  - scientific
  - technological
  - training and education
  - economic
  - social and societal

- Standard indicators (58 - 25+33)

- Includes:
  - name of indicator
  - detailed description
  - data needed

OECD. Reference framework for assessing the scientific and socio-economic impact of research infrastructures. (2019) doi:https://doi.org/10.1787/3ffee43b-en

Key performance Indicators (KPIs) to be used in the periodic review of ESFRI Landmarks

RACER (**R**elevant, **A**ccepted, **C**redible, **E**asy to monitor, **R**obust)

- **Relevant** – i.e. closely linked to the objectives of the RI over a particular period of time.

- **Accepted** by the RIs (at all levels) and stakeholders otherwise there will be limited implementation.

- **Credible** for non-experts, unambiguous and easy to interpret.

- **Easy** to monitor – e.g. data collection should be possible at low cost.

- **Robust** – e.g. against manipulation.

Report of the ESFRI Working Group on monitoring RIs performance | www.esfri.eu.
https://www.esfri.eu/latest-esfri-news/report-esfri-working-group-monitoring-ris-performance (2019)

Objectives relevant to RIs in general:

- Enabling scientific excellence
- Delivery of education and training
- Enhancing transnational collaboration in Europe
- Facilitating economic activity
- Outreach to the public
- Optimising data use
- Provision of scientific advice
- Facilitating International co-operation
- Optimising management

Report of the ESFRI Working Group on monitoring RIs performance | www.esfri.eu. https://www.esfri.eu/latest-esfri-news/report-esfri-working-group-monitoring-ris-performance (2019)

Quantitative indicators

Objective	KPIs
Enabling scientific excellence	1. Number of user requests for access 2. Number of users served 3. Number of publications 4. Percentage of top (10%) cited publications
Delivery of education and training	5. Number of master and PhD students using the RI 6. Training of people who are not RI staff
Enhancing collaboration in Europe	7. Number of members of the RI from ESFRI countries 8. Share of users and publications per ESFRI member country
Facilitating economic activities	9. Share of users associated with industry and publications with industry 10. Income from commercial activities and the number of entities paying for service
Outreach to the public	11. Engagement achieved by direct contact 12. Outreach through media 13. Outreach via the RI's own web and social media
Optimising data use	14. Number of publicly available data sets used externally
Provision of scientific advice	15. Participation by RIs in policy related activities 16. Citations in policy related publications
Facilitating international co-operation	17. Share of users and publications per non-ESFRI member country 18. International trainees 19. Number of members of the RI from non-ESFRI countries
Optimising management	20. Revenues 21. Extent of resources made available

Table 1. Numerical KPIs per objective. Further details of each KPI are provided in reference sheets **(Annex 5).**

Qualitative indicators

Annex 4. Examples of qualitative indicators per objective

Objective	Rational (what to measure)	Proposed Indicators*	Type
Enabling scientific excellence	Attractiveness of RI	5-year trend in number of proposals /user requests/registered users	Narrative
	Added value to science	Impact studies	Narrative
Enhancing collaboration in Europe (Sub-objective) Integration of distributed facilities	Policies related to integration of distributed RIs	A single access point to RI's data, services and/or facilities, as a prevailing more of access	Y/N
	Policies related to integration of distributed RIs	A single access point to resources of multiple partners of a distributed RI by industry	Y/N
	Policies related to integration of distributed RIs	Centralised evaluation and selection, based on excellence	Y/N
	Policies related to integration of distributed RIs	A common strategy and policy for intellectual property and know-how protection and exploitation adopted	Y/N
	Policies related to integration of distributed RIs	A central communication strategy adopted by the GA A joint research infrastructure roadmap	Y/N
	Policies related to integration of distributed RIs	A research infrastructure roadmap of the RI	Y/N
		Existence of an Industry Engagement Plan	

...

Data Sheets for KPIs

A. ENABLING SCIENTIFIC EXCELLENCE

1. Number of user requests for access

Objective	Enabling Scientific Excellence
Indicator	Number of user requests for access
Definition(s)	For access to facilities: number of user proposals for access. For resource RIs: number of users of resources, such as collections, data, services
Rationale	Indicator of the attractiveness of the RI.
Assumptions	The size of the community depends on a number of factors. Young, relatively unknown RIs will start with a small community which will increase with their visibility, quality and extent of offer. In the case of the requests for access to the research facilities, the success rate may affect the number of applications. Once it reaches a very low level, the number of applications is likely to level off. In some RIs, metrics on usage requests is maintained at the level of individual resources, as researchers don't make central requests for access. In the case of resource RIs, the number of users is also affected by the terms of access. Some RIs require no registration, while others demand registration processes of varying complexity. Also, in the case of monitoring of access though IP address it needs to be considered that in some cases, hundreds of users may use the resources though one IP address.
Data/information needs and resources	A tracking/recording system should be set up by the RI.
Who is providing this information	Provided by the RI.
Detailed methodology for indicator calculation	Record and report the number of applications for access/registered users. For RIs which do not require registration, number of unique users/visits/logins is reported. For RIs providing more than one type of service (e.g. data, services, access to the facility, platform and event-based access), values for each category are reported. Subgroups may be reported, as per - Share of users per ESFRI country (KPI 8); - International users (KPI 18); - Academic users; - Non-proprietary Industrial users (KPI 9).
Unit of measure	Number
Frequency of measurement	Annually
Assessment of indicator quality and comparability	Commonly used.
Estimated cost of data collection (including access to external databases)	Generally low. May be high for RIs offering fully open and free access to resources.
Level of reporting burden	Generally low. May be high for RIs offering fully open and free access to resources.
Additional issues or Observations	International: non-ESFRI member countries. Some RIs may not distinguish between KPI 1 and 2 and will report either of the two. In such a case, the assumptions described here may apply to KPI 2. Technical solutions may improve the quality and ease of reporting.

Items

- Name of the Indicator
- Definition
- Rationale
- Assumptions
- Data/info needed and resources
- who is providing the information
- detailed methodology
- unit of measure
- frequency
- indicator quality and comparability
- cost of data collection
- level of reporting burden
- other issues/observations

Example:

ACTRIS – The Aerosol, Clouds and Trace Gases Research Infrastructure

Impact category	KPIs
Human capital creation	number of PhD students number of MSc students number of graduates trained number of foreign students trained research and technical staff attracted to be employed % of research and technical staff attracted from abroad
Scientific activity	number of articles published in journals number of articles published in proceedings number of books (chapters) number of PhD dissertations number of patents' applications number of patents granted number of methodologies, prototypes or designs number of scientific events

Impact category	KPIs
Innovation	total volume of funding to R&D projects number of start-ups and spin-offs total turnover and earnings from start-ups and spin-offs procurement contracts number of companies in development and upgrade of equipment
Macroeconomic effects	total spending for developing and operating total impact on gross value added from devel and operation total impact on available income attributed to devel and operation impact of gross value added attributed to research direct, indirect and induced effects on employment attributed to research total impact on available income

Socio-economic impact - challenge
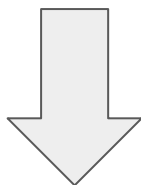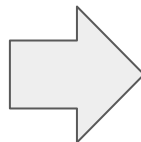
Word-cloud from DISSCO proposal - socio-economic impact
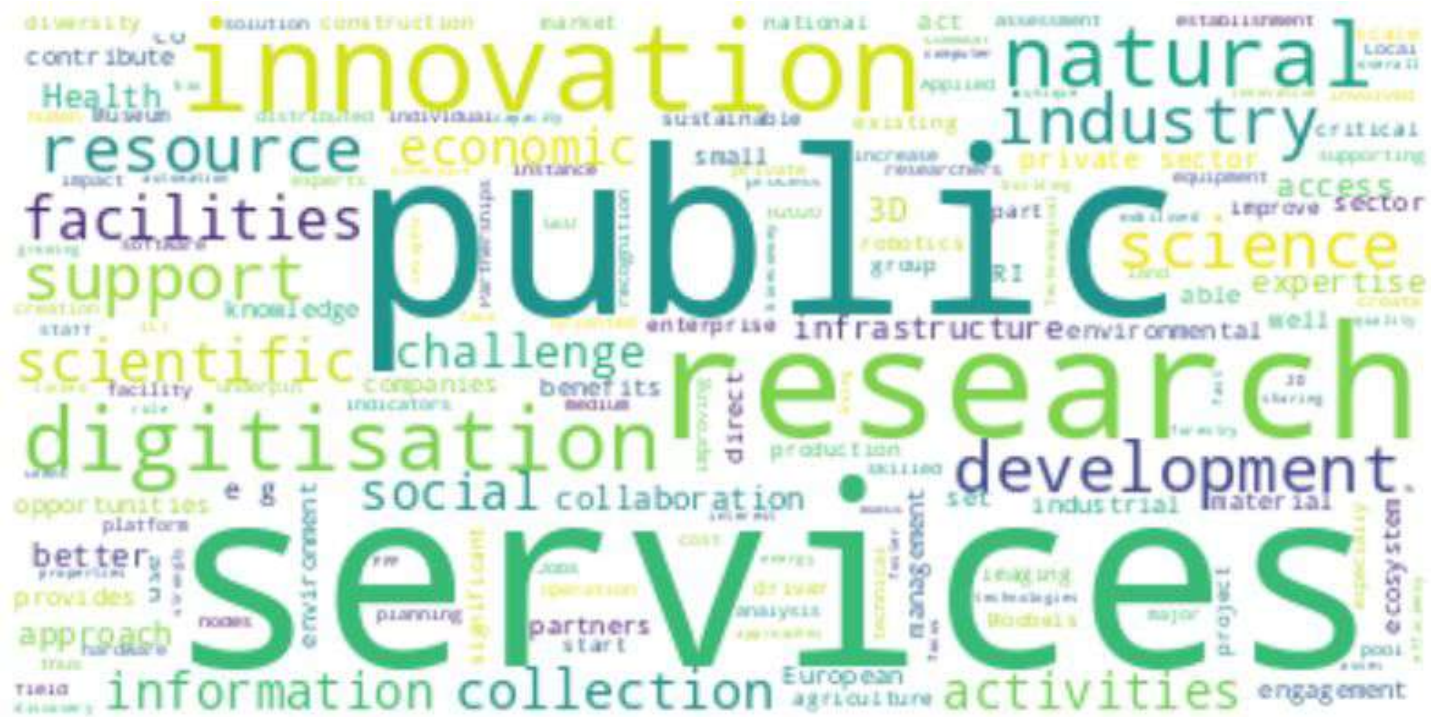
DISSCO proposal - socio-economic impact

Scientific
- DNA barcodes, genomes, proteomes and metablomes
- 2D/3D imaging

Industry and innovation
- information science (big data)
- computer vision
- 2D/3D scanning
- new pharmaceuticals (combining collection data with metablomic
- new cultivars and animal breeds
- new standards
- new materials inspired by nature

DISSCO proposal - socio-economic impact

Direct socio-economic impacts
- job creation
- industry-oriented economic benefits
    - impact on organisations
    - applications in agriculture, environmental assessment, land use planning
    - new hardware/software - SME's

Mid and long term socio-economic benefits
- Economy of scale
    - common digital data processing
    - purchasing equipment

- Economy of scope
    - industrialization of digitisation
    - robotics, optics, imaging

DISSCO proposal - socio-economic impact

Innovation activity in the production of goods and services
- Direct contributions to food, textile, building materiais, medicines, provision of sustainable energy, rare minerais, ecosystem services

Role in:
- Technological innovation - critical step for its implementation, direct spin-off, driver for industry-led innovations
    - (meta-)data standardisation, information management, computer vision, robotics and automation, and 2D/3D imaging

- Social innovation - citizen science and crowdsourcing focus (through the museum's traditional focus) in public engagement

DISSCO proposal - socio-economic impact

Attract innovation-oriented resources

- **Industry as supplier** - 2D/3D imaging, robotics and automation, image/pattern recognition algorithms as well as information management technologies

- **Industry as user** - companies will be able to augment their datasets with quality information on the natural world

Tackling (grand) societal challenges:
- DiSSCo data and expertise can directly contribute to "ecosystem health"
- genetic material support the development of new agricultural varieties
- describing and understanding bio- and geo-diversity on earth

SOCIO-ECONOMIC IMPACT INDICATORS

REQUIREMENTS FOR THE CREATION AND OPERATIONALIZATION OF SOCIO-ECONOMIC IMPACT

E. FONTAINHA

JANUARY 18-22 2021

USER NEEDS & **SOCIOECONOMIC IMPACT** (WP1)

- *User needs & socioeconomic impact (WP1) Contributes to Scientific Readiness objectives*

- *Building on surveys and networking activities already carried out in previous projects and networking activities, this WP will examine the needs of research projects and research application activities (e.g. nature conservation) based on NSCs, and the requirements these needs set for the services to be provided by DiSSCo. This will be done across biological (Task 1.1) and geological (Task 1.2) NSCs.*

- *Criteria for prioritization of digitization of collections, data linking, generation and enrichment, taking into consideration user needs and (potential) socioeconomic impact, will be analyzed (Task 1.3).*

- **Socioeconomic impacts of DiSSCo and DiSSCo-enabled research will be examined with the aim to develop indicators and tracking mechanisms (Task 1.4).**

- ***"the answers to the 'questionnaire', covering issues related to the action implementation and the economic and societal impact, notably in the context of the Horizon 2020 key performance indicators and the Horizon 2020 monitoring requirements"*; (GA p. 32) (content of Periodic Report)**

SOCIO ECONOMIC IMPACTS

- long-term effects produced by a development intervention

- Positive and negative,

- primary and secondary,

- directly or indirectly

- intended or unintended.

- …

INDICATORS IN DISSCO PREPARE (EXAMPLES)

- *WP 1 1.4 Develop indicators of socioeconomic impact*

- ***WP 3 3.2.4 Digitisation Monitoring****: Ongoing and prospective digitisation activities will be highly distributed across the DiSSCo consortium. To monitor digitisation pipelines, a set of digitisation measures and dashboards will be needed.* **This subtask will put in place the framework for these, supporting delivery of a set of agreed key performance indicators (KPIs) for digitisation.** *This will allow the progress of digitisation to be monitored across the consortium (p.17)*

- **WP 4 Subtask 4.1.1 Develop a set of indicators for estimating the cost** *of running the infrastructure and providing services.*

- **WP 6 Subtask 6.4.4.** *Enable FAIR data and compliant services, and integration in the EOSC service catalogue This subtask will develop and adopt* **measurements and indicators** *in collaboration with ENVRI-FAIR and the Go- Fair initiative* **to measure compliance of data and services with the FAIR data principles.**
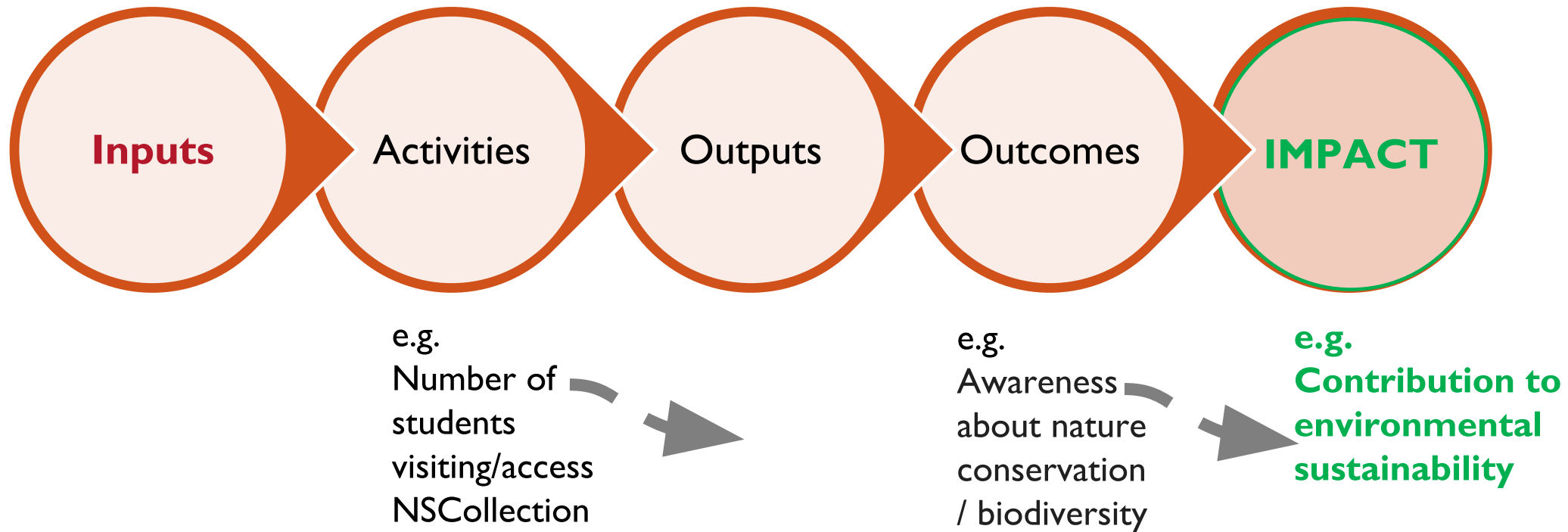
QUESTIONS AND CHALLENGES

- Will the indicators used/created in DiSSCo Prepare (performance, impact, etc.) be combined in a single system of indicators?

- Will the indicators (selection, etc.) used/created in DiSSCo Prepare (performance, impact, etc.) be discussed among the different Task Teams?

- How are User Cases and User Stories (WP 1.1. and WP 1.2.) and (WP 1.3. and WP 1.4.) connected ? The results from WP 1.1. and WP 1.2. provide inputs (main? some?) to WP 1.4. ?

- Is the Criteria Prioritization Model for Digitalization (WP 1.3.) combined with Develop Indicators of Socioeconomic Impact (WP 1.4.)? Or, alternatively, are they developed independently (see ECIDIG.EU survey results)

- The RACER (Relevant, Acceptable, Credible, Easy and Robust) criteria are suitable (sufficient, etc.) to apply in all domains (in all RI)?

- Because RIs are different the indicators systems include 'core indicators' (common to all RI s) and additional indicators. Self tailored systems.

CHALLENGES & PRIORITIES

- *"It remains challenging to elaborate **a robust model and metrology to quantitatively estimate the Socio-Economic Impact (SEI)**."*

  http://roadmap2018.esfri.eu/landscape-analysis/section-3/socio-economic-impact/

- *"A separate question asked respondents to indicate which of the **4 relevance categories were most relevant** when assessing digitization priorities. This ranking allowed us to discover which prioritization criteria received the highest value."* **Innovation and consolidation for large scale digitisation of natural heritage.** ICEDIG.EU Deliverable 2.1. p. 5.

- 4 Categories. What is the most relevant?

  - Collection ?

  - Economic ?

  - Scientific ?

  - Social ?

"WHAT IS ACCORDING TO YOU THE MOST RELEVANT CATEGORY FOR PRIORITIZATION OF COLLECTION DIGITIZATION?"
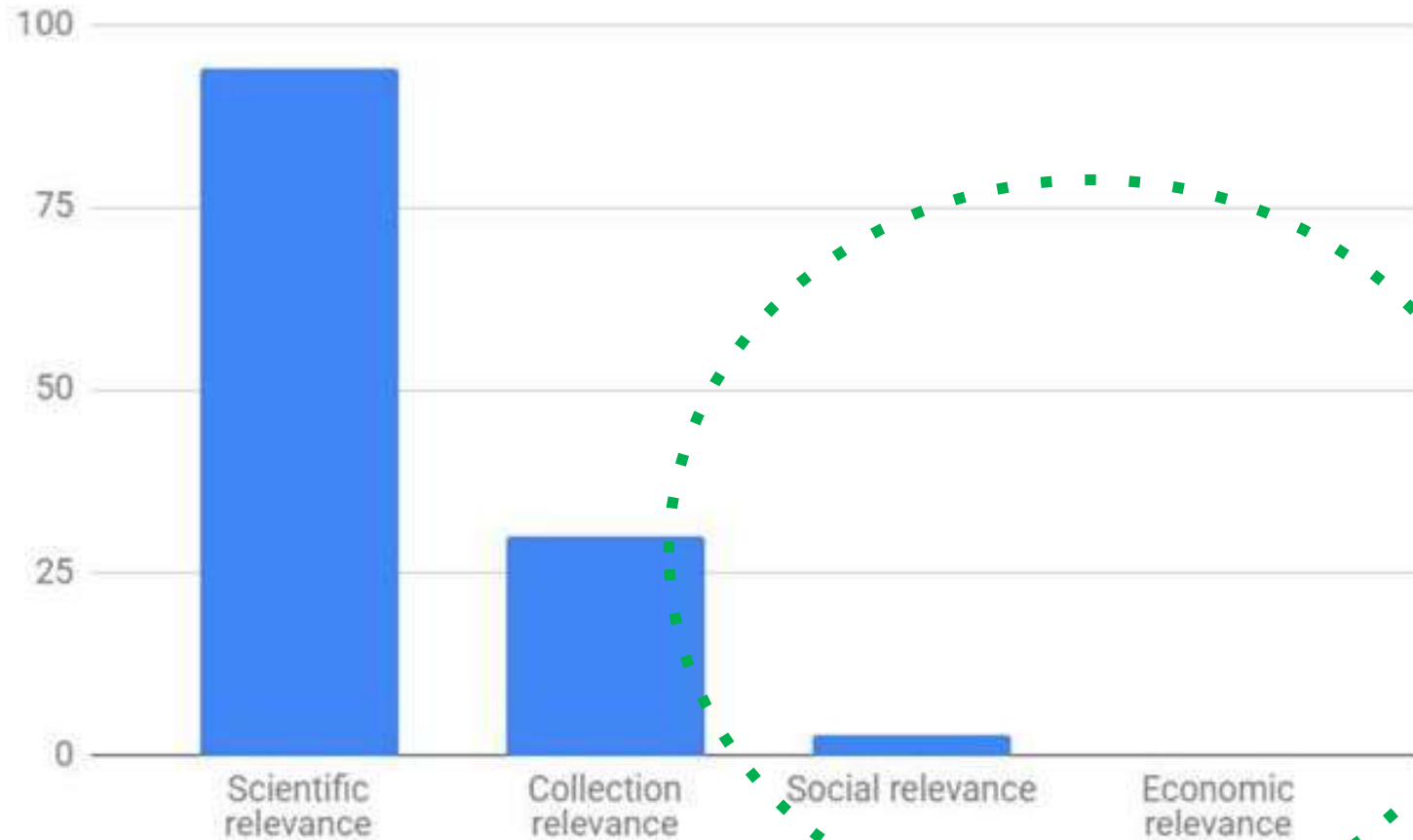
Stakeholders Preferences

Fig. 15. Relevance categories (Q16): What is according to you the most relevant category for prioritization of collection digitization? From left to right: Scientific relevance (94), Collection relevance (30), Social relevance (3) and Economical relevance (0).

9

RI PROJECT EVALUATION IMPACT (EXAMPLES)

Project	Impact Area
RI-PATHS	ECONOMY AND INNOVATION Added value of RI-owned patents and other IP Corporate efficiency gains through use/application of RI data Technological impact: Number of new technologies and designs Market creation impact: triggered sales volume Market expansion impact: increased sales volume Market expansion impact: increased revenues Increased economic activity in the region/nation
	POLICY Notable changes in relevant regulations Notable changes in funding decisions Increased trust in science Notable changes in policy decisions

RI PROJECT EVALUATION IMPACT (EXAMPLES)

Project	Impact Area
RI-PATHS	SOCIETY Contribution to public sector challenges: Administration & governance Contribution to social sustainability: CSR, Social Inclusion, Culture Contribution to Gender balance Contribution to environmental sustainability: Energy & Waste issues Improvement of wellbeing: Health & Ageing Inclusion of topics in schools and academic curricula
	HUMAN RESOURCES Scientific attractiveness Improvement of HRST (C) in region/country (Scientific) Improvement of HRST (C) in region/country (Technical/Managerial) Improved job opportunities in the region/nation Increased Prestige as Training Facility

MEASURING THE IMPACT

WHICH INDICATORS CAN WE BE INNOVATIVE, RACER, ETC.

health

Agriculture food and food security

nature conservation

education

citizen science

other

Agenda

Time	Topic	Lead
10.45-10.50	Welcome	AC
10.50-10.55	1)Aims	AC
10.55-11.25	2) T2.1 Training Strategy: Task Presentation	JA
11.05-11.20	3) Alignment with sister initiatives & previous development	AC/HH/HdB /CP/MC
11.20-11.45	Tour de table	All
11.45-12.05	4) T2.1 Work plan draft: presentation	MLK
12.05-12.10	5) Identify next steps	MLK
12.10-12.15	AOB	AC

1) AIMS OF THE MEETING

Kick-off Meeting

1) Common **understanding** of the task

2) Establishment of a regular **working channel**

3) **Distribution of work** and responsibilities in accordance to partners' expertise.

T2.1 – KoM
1) Common understanding

Partners and their contribution

CETAF (5), NHM-UIO (4), ULISBOA (3), MfN (3), UniFi (2), MNHN (1)

Duration

February 2021 until December 2022 (M13 - M35)

Objective

Deliverable D2.1: Training Strategy report
Documenting the **key components** on which to build the DiSSCo training strategy and drafting **recommendations** on how to tackle and further integrate all of those features.

MS2.1: Recommendations on suitable training mechanisms (M23)
MS2.2: Landscape analysis of BPs for training delivery completed (M28)

T2.1 – KoM
1) Common understanding

Timeline: (23 months)

Kick-off Meeting	1st Milestone	2nd Milestone	DELIVERABLE D2.1
22 January 2021	December 2021 (M23)	Mayo 2022 (M28)	December 2022 (M35)

**Definition of scope
Work Distribution
Communication channels**

Recommendations on **training mechanisms (MS2.1)**

Analysis of BPs (MS2.2)

Training strategy report: BPs and recommendations **(D2.1)**
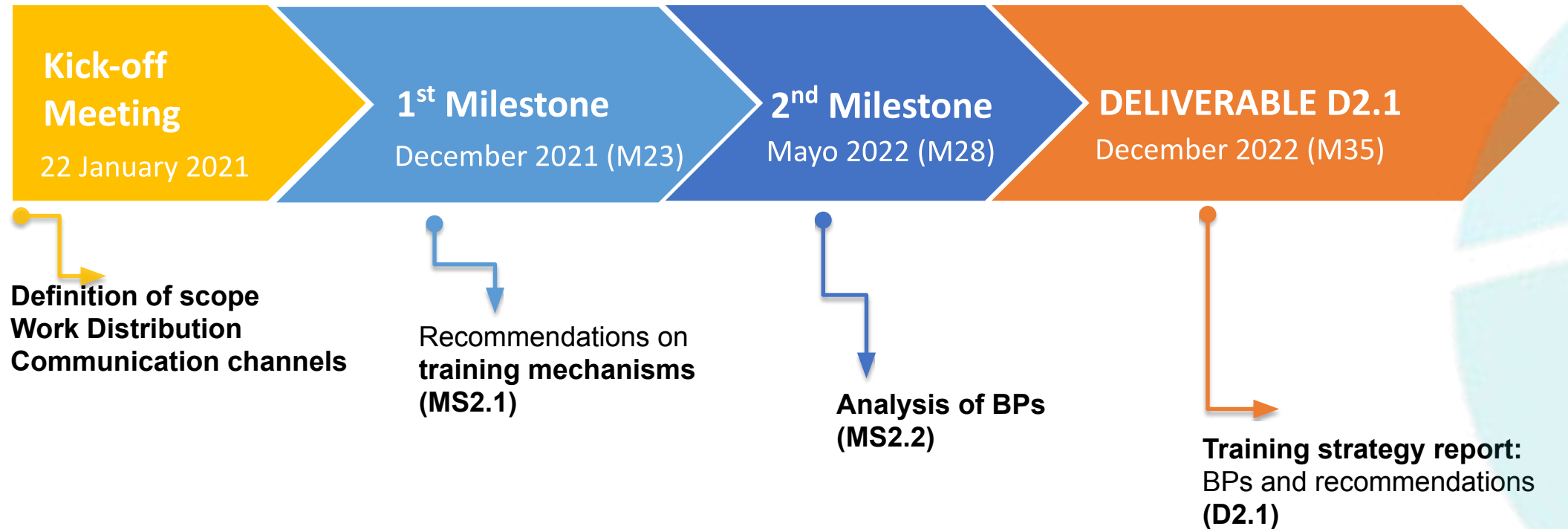
T2.1 – KoM
1) Common understanding

Description:

Business model, trainers, maintenance & sustainability, resources allocation

Platforms, portals

(…) to develop a training strategy with distinct **channels** and **modes of accessing** training to address the identified **needs**. Personnel capacity building will consider both the side of the **data suppliers** and the side of the **users of DiSSCo.**

Collections CMS, standards, protocols and workflows, APIs

Scientists, researchers in the domain and beyond, other Ris, CS, Policy makers

Material available, gaps, academic & professional training, diversity, BPs
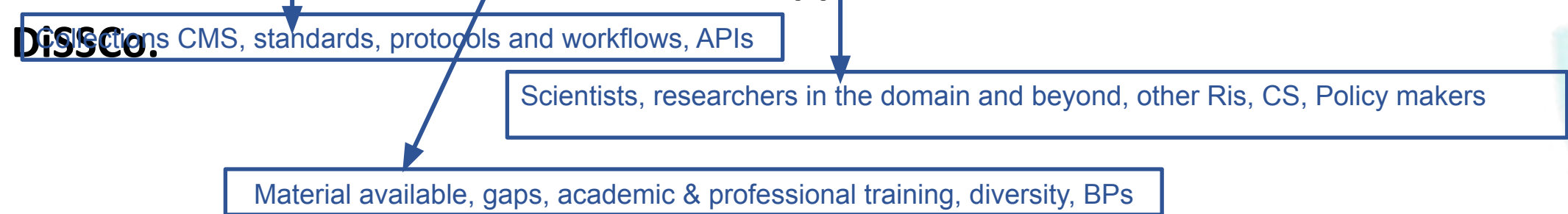
T2.1 – KoM

1) Common understanding

Sub-tasks:

a) compilation of needs for skills/competencies;
b) identification of training providers/platforms;
c) landscape analysis of best practices for training delivery; and
d) integration of all training strategy elements,

to jointly provide a final report with **recommendations** for setting up the DiSSCo Training Strategy.

T2.1 – KoM

1) Common understanding

Supporting resources:

ICEDIG Design Blueprint – Recommendations R74 to R100
SYNTHESYS+ D2.3: Schema for the training programme (Apr'21)
DiSSCo DPP MS3.1: Consultation and analysis (Dec'20)
DiSSCo DPP D3.1: Digital skills and competencies dashboards (Ag'21)

Work plan definition:
To be worked out collaboratively

T2.1 – KoM

2) Communication channels

Responsible people:

ULisboa: Judite Alves, Maria Joao Ferreira Nicolau dos Santos

NHMUO: Hugo de Boer and Dag Endresen

MfN: Jana Hoffmann and Christoph Häusser

UniFI: Luca Bartolozzi, Gianna Innocenti

MNHN: Anne-Sophie and Archambeau and Jonathan Blettery

Task meetings: bi-weekly,1hr max

Repository of documents: TW

T2.1 – KoM
3) Distribution of work

2) CONTEXT FOR T2.1

WP2. Human resources, Training & Users Support

Overarching mission:

Ensure that the DiSSCo community has the appropriate skills and competencies to use and manage the infrastructure efficiently.

WP2. Human resources, Training & Users Support

Key objectives:

To create a **training strategy** that addresses the identified needs for both data suppliers and end-users.

To design **user support services** oriented towards offering information and supporting to the user in need to use of the DiSSCo services and facilities.

To develop a **Human Resources Policy** that ensures that RI-related human resources have the appropriate skills and abilities and are used effectively and efficiently to accomplish DiSSCo objectives.

Each objective is addressed by an individual task

To create a **training strategy** that addresses the identified needs for both data suppliers and end-users. ▶ **T2.1. Training Strategy**

To design **user support services** oriented towards offering information and supporting to the user in need to use of the DiSSCo services and facilities.

T2.2. Helpdesk and User Support Services

To develop a **Human Resources Policy** that ensures that RI-related human resources have the appropriate skills and abilities and are used effectively and efficiently to accomplish DiSSCo objectives. ▶ **T2.3. Human Resources Policy**

WP2.

Training strategy

Training for HD use

HD as main channel

DiSSCo Community
data providers + end-users

Framing of training models in HRP

Helpdesk: Support services

Human Resources Policy

Interactions with other DPP WPs/Tasks and other DiSSCo linked projects

DiSSCo linked-projects
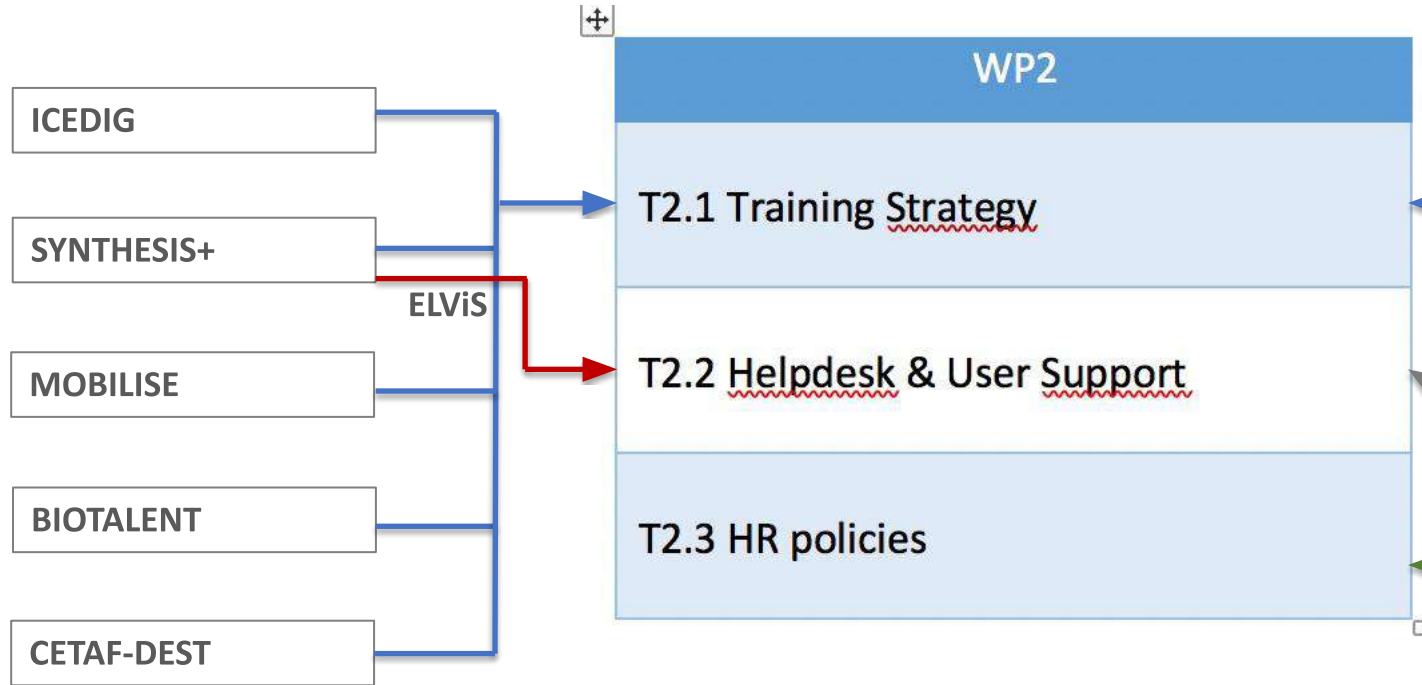
DiSSCo Prepare WPs/Tasks

ICEDIG

SYNTHESIS+

ELViS

MOBILISE

BIOTALENT

CETAF-DEST

WP2

T2.1 Training Strategy

T2.2 Helpdesk & User Support

T2.3 HR policies

T3.1. Digital skills across facilities

T3.3. Develop DiSSCo secondment and distributed team working practices

T7.2. Legal Entity

WP6. Technical Architecture & Service Provision
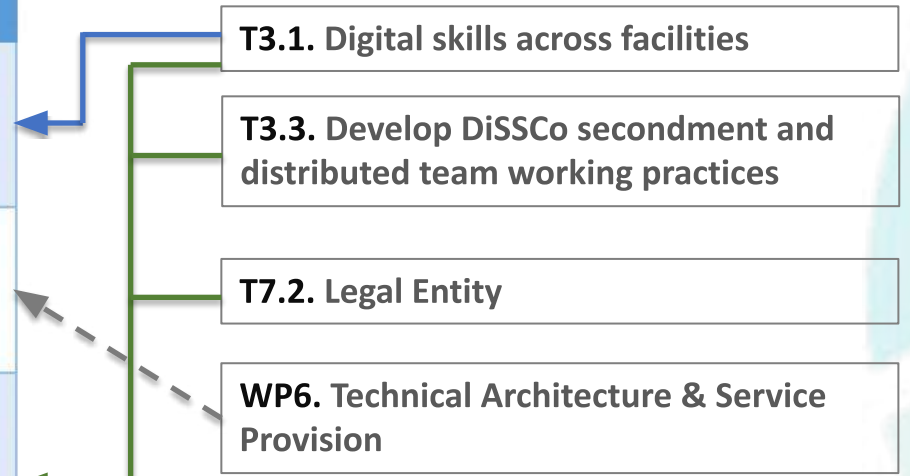
Any questions?

3) ALIGNMENT

Sister initiatives & previous developments

- **DiSSCo PPP WP3**: T3.1 "Improve digital skills and competencies across DiSSCo facilities" *by Helen Hardy (NHM London)*

- **DEST: Distributes European School of Taxonomy** *by Hugo de Boer (UO_NHM Oslo)*

- **SYNTHESYS+ T2.3** *by Carole Paleco (RBINS Brussels) & Magalie Castelin (MNHN Paris)*

Tour de Table

- Affiliation
- Expertise
- Possible contribution

4) T2.1 - WORK PLAN DRAFT

Subtasks according to task outcomes

MS2.1 Recommendations on suitable training mechanisms

Subtask 1 - Compilation of needs for skills/competences
Subtask 2 - Identification of training platforms and providers

MS2.2 Landscape analysis of BPs for training delivery completed

Subtask 3: landscape analysis of best practices for training delivery.

D2.1 Training Strategy

Subtask 4 - Integration of all training strategy elements, to jointly provide a final report with a recommendation for setting up the DiSSCo training.

2021

January - Work plan presentation during AHM1

TOWARDS MS2.1: February to November

Subtask 1 - Compilation of needs for skills/competences (Jan-July) *(1)*

- **Prepare the scenario**: actions to be implemented ahead of Synthesys+ D2.3 (due April 2021):
  - *REVIEW* ICEDIG recommendations
  - *DEFINE* the components to be considered in a training programme
  - *DISTRIBUTE* the work among partners.
  - *COLLECT* recommendations / resources from sister initiatives or other infrastructures.

2021

January - Work plan presentation during AHM1

TOWARDS MS2.1: February to November

Subtask 1 - Compilation of needs for skills/competences (Jan-July) *(2)*

- Agree on what the **training mechanisms** are/could be.

- Review the **key takeaways from SYNTHESYS+ D2.3**: schedule working meeting with task partners in May 2021.

- Define **how to implement recommendations from D2.3** in regards to training mechanisms.

2021

TOWARDS MS2.1: February to November

Subtask 2 - Identification of training platforms and providers (Aug-Dec)

- **Set the structure** of the training programme

- **Identify** the available platforms

- **Establish a methodology** (criteria & requirements) to link platform programme content material to the suitable & select providers.

- **Draft** the recommendations *(November)*

- **Review** the draft report of recommendations *(December)*

2022

January - Submit MS2.1

TOWARDS MS2.2: January - June

Subtask 3: landscape analysis of best practices for training delivery.

- *June:* Submit MS2.2 Landscape analysis of BPs for training delivery completed

TOWARDS D2.1: July - December

Subtask 4: Integration of all training strategy elements, to jointly provide a final report with a recommendation for setting up the DiSSCo training.

- *December:* Draft of D2.1 Training Strategy Report.

2023

January - Submit D2.1

Open discussion Feedback on the draft

5) NEXT STEPS

- Provide **feedback** on the [work plan draft](deadline: 29 Jan) (deadline: 29 Jan)

- Schedule **frequent meetings**:
  - Define frequency and timeslot

- Set up **online work mechanisms** to collect information from DiSSCo partners.
  - Teamwork tools
  - Google suite.

AOB

"Bringing the irreplaceable data stored in natural science collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

The preparatory phase project of DiSSCo Research Infrastructure - Distributed System of Scientific Collections

https://www.dissco.eu/prepare/
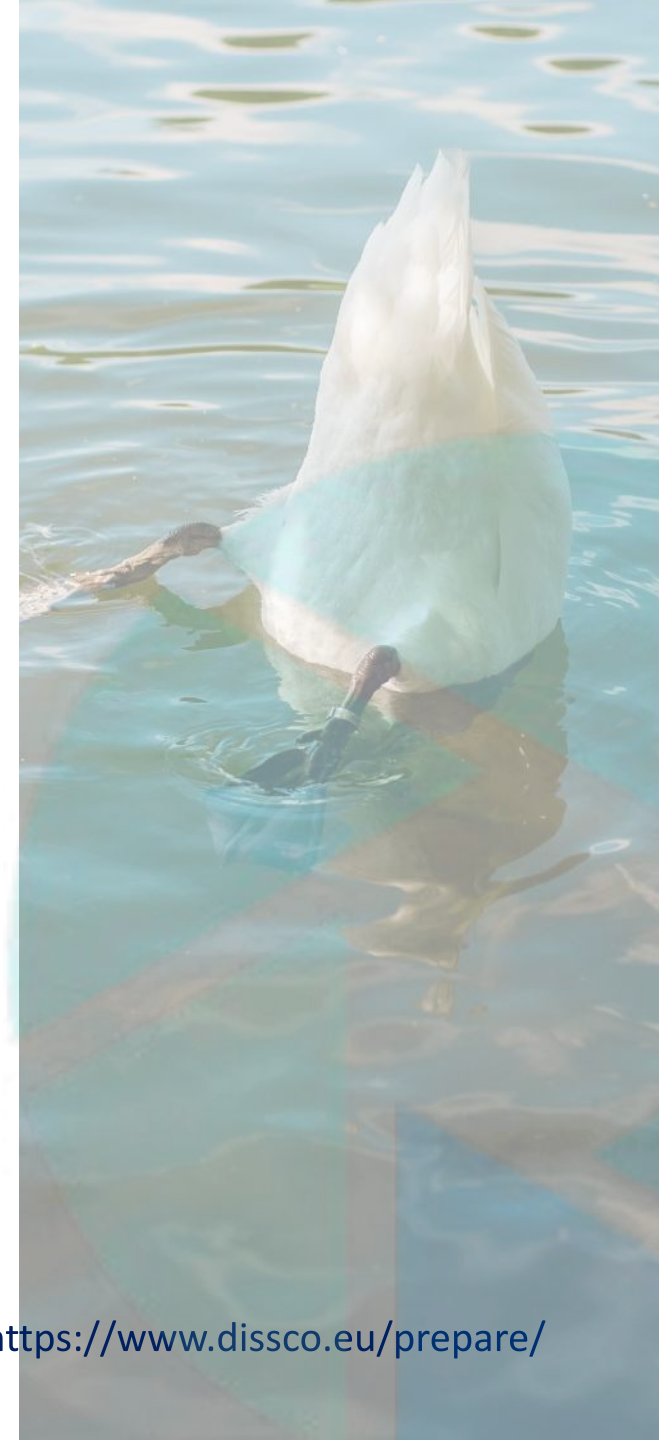
All Hands Meeting1 T2.1 Training Strategy session

Alignment with sister initiatives

SYNTHESYS+ *NA 2.3: Develop focused trainings activities*

Task co-leaders: Magalie Castelin (MNHN)

Carole Paleco (RBINS)

Network Activities Task 2.3: Develop focused training activities

- 13 Partners, 32 persons, 11 Institutions **+** CETAF-DEST, GBIF, TDWG.

- Started **SYNTHESYS+ Kick off Meeting** (London, 02/2019).

- A single deliverable, **D2.3 (Catalogue of Training modules)**, due after 26 months (04/2021).

- One milestone **MS26 (Workshop on Training Programme Definition)**, due after 12 months (02/2020).

NA Task 2.3: Develop focused training activities - Main Objectives

➢ Tackle both **TA and VA needs** by supporting the community in **acquiring digital (data) skills and competencies** that enable individuals/ researchers/curators to **navigate collections information effectively**, as well as **optimise access and utilisation** of NH collections.

➢ **Identify Training mechanisms** critically needed.

➢ Develop a set of specific training modules for delivery through face-to-face meetings and as online resources.

➢ **Targeted activities**, including train-the-trainers events, will focus on specific items such as supporting use of software packages for data access and analys**is.**

➢ **Several key training areas** - covering both VA and TA - can be foreseen, includings training on:

- Data mobilisation, processing, analysis, quality and access;
- data standards and interoperability;
- crowdsourcing and citizen science;
- bio- and geodiversity policy and legislation;
- multispectral digitisation and 2D+/3D digitisation;
- DNA barcoding and metabarcoding;
- use of ELViS platform;
- and software for collections management

Collaborative work process M1-M12

SYNTHESYS+ Kickoff Meeting (London, Feb 2019).
T2.3 Partners meet and agree on the need to align with **DiSSCo Training Needs** and to Focus on trainings expanding **digital workforce skills in big data processing.**

Explore and gather data about existing training modules & e-tools across T2.3 Institutions, regardless of how each training would be prioritised later on.

Understand and define **DiSSCo Training Needs**.
Planning the T2.3 Work Plan for 2019-2020.
Expend collaborations with DEST and stakeholders.
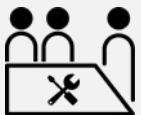
Workshop on Training Programme definition MS26 (Paris).
(1) Introduction of existing training modules and tools;
(2) Definition of Key Training Areas (KTAs);
(3) Identification of gaps and settling discussion about how to fill these gaps;
(4) Seek agreement on prioritisation of training topics.

Outcomes of the workshop (Jan 22-23, 2020 – MNHN Paris)

NA2.3 Advert. Poster

NA2.3 MS26 Report, 15pp.

NA2.3 Catalogue of Training Modules.

Outcomes of the workshop (Jan 22-23, 2020 – MNHN Paris)

NA2.3 Catalogue of Training Modules.

Data collected

- Institution, City
- Audience
- Length
- Frequency
- Physical, virtual, or hybrid
- Certificate
- Fees

Key Training Areas (KTAs)

- Musuem specimen
- Data
- Policy & Legislation
- Citizen science
- …

Training gaps

- Collection assessment
- Use of ELViS
- Data standards
- Citizen science
- ABS
- Open Science Policy

Analysis *in progress*
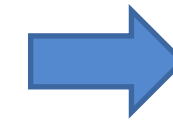
Description of the training landscape & current training mechanism.

Launch of DiSSCo PP (Feb 2020) - New objectives & scope
How to go further into:
- Expending the Catalogue.
- Refining gaps **and Training Needs Assessment.**
- Providing recommendations for the development of proactive, efficient, and evolving
DiSSCo Training Programme (DPP-WP2).

New Analysis: IRLs *vs* Trainings (M15)
(1) Science
(2) Data
(3) Technological
(4) Organisational
(5) Financial

Deliverable 2.3 (M24)
Disseminate catalogue **(DEST)**
Roadmap to the Development of DiSSCo Training programme
Workflow: adapting the Training Programme to the constant evolution of the training needs

DiSSCo training needs and gap analysis

IRLs DiSSCo Prepare	Matching to training
WPs and definition	**Sub-categories**
SCIENCE <u>User needs & socioeconomic impact (WP1)</u> <u>Human Resources, Training & Users Support (WP2)</u> DiSSCo, like any RI, needs to be in full and constant alignment with the needs of its identified user base (see above). This requires the existence of a practical scientific evaluation framework, which allows the RI to inform decisions around its future scientific programme based on the ever-changing needs of its user base.	Collection/sampling of specimen Specimen curation Museum specimen based research Taxonomy Ecology Biodiversity Geodiversity Evolution Data analysis Publication and outputs Citizens science - training the trainers Citizens science - training citizens
DATA <u>Capacity enhancement (WP3)</u> As a distributed data infrastructure, DiSSCo relies on the capacity of the national and institutional nodes to provide and enrich data in a consistent, harmonised manner, compatible within the overall RI.	Pre-digitisation curation Standarisation interoperability Data curation Digitisation Databasing Data sharing
TECHNOLOGICAL <u>Common Resources and Standards (WP5)</u> <u>Technical Architecture & Services provision (WP6)</u> Predominantly operating as a data infrastructure, DiSSCo needs to fully understand the current limitations of technological implementations of existing e-infrastructures and subsequently establish its data architecture as	Operating Develonement IT data managment e-Monitoring
ORGANISATIONAL <u>Governance, Policy & Legal frameworks (WP7)</u> To proceed with the set-up of its new legal entity, DiSSCo needs to further set out its overall organisational (governance and management) structures, strategic and operational plan, the involvement of funders and infrastructure siting as well as proceed with harmonisation of policies	Policy Governance Management Communication
FINANCIAL <u>Business Framework (WP4)</u> Before DiSSCo fully embarks on implementation activities, it is necessary to have a robust financial framework in place that enables accurate calculations of costs and contributions. This framework should deliver the	Management Financial awarness Basic information

- **Objectives & scope of DPP WPs**
- **Define sub-categories using DPP WP4 Cost Book**

  - The *Cost Book* is a comprehensive list of the **skills** and **tasks** needed to run a **Research Infrastructure** willing to embark on the **big-data era**, **massive digitization-on-demand**, and high throughput production of **scientific knowledge based on natural data,** with all the **government's policies** that it does involve.

IT data management: all activities surrounding the manipulation, protection, storage, or safety of data.

What are the potential training needs of each phase for a self-sustainable RI?

Proposal Phase 2015 - 2018	Preparatory Phase 2018 - 2021	Construction Phase 2020 - 2023	Transition Phase 2023-2024	Operation 2025 – 2035

DiSSCo training needs and gap analysis

IRLs DiSSCo Prepare	Matching to training
WPs and definition	**Sub-categories**
SCIENCE User needs & socioeconomic impact (WP1) Human Resources, Training & Users Support (WP2) DiSSCo, like any RI, needs to be in full and constant alignment with the needs of its identified user base (see above). This requires the existence of a practical scientific evaluation framework, which allows the RI to inform decisions around its future scientific programme based on the ever-changing needs of its user base.	Collection/sampling of specimen Specimen curation Museum specimen based research Taxonomy Ecology Biodiversity Geodiversity Evolution Data analysis Publication and outputs Citizens science - training the trainers Citizens science - training citizens
DATA Capacity enhancement (WP3) As a distributed data infrastructure, DiSSCo relies on the capacity of the national and institutional nodes to provide and enrich data in a consistent, harmonised manner, compatible within the overall RI.	Pre-digitisation curation Standarisation interoperability Data curation Digitisation Databasing Data sharing
TECHNOLOGICAL Common Resources and Standards (WP5) Technical Architecture & Services provision (WP6) Predominantly operating as a data infrastructure, DiSSCo needs to fully understand the current limitations of technological implementations of existing e-infrastructures and subsequently establish its data architecture as	Operating Developement IT data managment e-Monitoring
ORGANISATIONAL Governance, Policy & Legal frameworks (WP7) To proceed with the set-up of its new legal entity, DiSSCo needs to further set out its overall organisational (governance and management) structures, strategic and operational plan, the involvement of funders and infrastructure siting as well as proceed with harmonisation of policies	Policy Governance Management Communication
FINANCIAL Business Framework (WP4) Before DiSSCo fully embarks on implementation activities, it is necessary to have a robust financial framework in place that enables accurate calculations of costs and contributions. This framework should deliver the	Management Financial awarness Basic information

- **Objectives & scope of DPP WPs**
- **Define sub-categories using DPP WP4 Cost Book**

  - The *Cost Book* is a comprehensive list of the **skills** and **tasks** needed to run a **Research Infrastructure** willing to embark on the **big-data era**, **massive digitization-on-demand**, and high throughput production of **scientific knowledge based on natural data,** with all the **government's policies** that it does involve.

Development: all the skills necessary to design databases, software, and digital tools (e.g., data architecture, data carpentry – i.e., development of ELViS).

What are the potential training needs of each phase for a self-sustainable RI?

Proposal Phase 2015 - 2018	Preparatory Phase 2018 - 2021	Construction Phase 2020 - 2023	Transition Phase 2023-2024	Operation 2025 – 2035

DiSSCo training needs and gap analysis

Training modules of the Catalogue

IRLs sub-categories		MNHN	RBINS	MNCN-CSIC	NHMW	UGOT	RBG Kew	RMCA	MfN	MfN/SNSB	NRM	TOTAL
Nb total trainings		7	4	10	18	2	6	6	11	3	4	71
SCIENCE	Collection/sampling of specimen	0	1	3	5	0	4	3	0	0	4	20
	Specimen curation	0	2	6	6	0	5	3	2	1	4	29
	Museum specimen based research	6	0	7	13	0	5	3	0	0	4	38
	Taxonomy	2	1	5	5	1	2	3	0	0	4	23
	Ecology	1	0	1	6	2	0	4	0	0	0	14
	Biodiversity	3	0	5	8	2	2	3	0	0	3	26
	Geodiversity	0	0	1	2	0	0	3	0	0	2	8
	Evolution	1	0	2	8	1	1	1	0	0	0	14
	Data analysis	5	0	1	9	2	3	5	0	0	0	25
	Citizens science - training the trainers	0	1	2	9	0	0	0	0	0	0	12
	Citizens science - training citizens	0	0	0	9	0	0	0	0	0	0	9
	Publication and outputs	1	0	2	6	2	2	4	1	0	1	19
DATA	Pre-Digitisation curation	1	1	4	1	0	4	2	1	0	3	17
	Digitisation	3	2	2	4	0	4	1	1	0	4	21
	Standarisation interoperability	3	1	2	2	0	4	2	4	0	0	18
	Data curation	1	0	4	2	2	5	2	4	0	1	21
	Databasing	1	1	7	1	0	5	6	4	0	1	26
	Data sharing	0	0	2	2	0	1	5	2	5	0	17
TECHNOLOGICAL	Operating	2	2	3	0	0	1	0	1	0	0	9
	Development	2	0	0	0	0	0	1	0	0	0	3
	IT data management	3	0	1	0	0	0	2	4	0	0	10
	e-Monitoring	0	0	0	0	0	0	0	0	0	0	0
ORGANISATIONAL	Policy	0	1	4	1	0	4	0	0	2	0	12
	Governance	0	0	0	1	0	0	0	0	0	0	1
	Management	0	0	0	1	0	2	0	0	0	0	3
	Communication	0	0	0	6	0	2	0	0	0	0	8
FINANCIAL	Management	0	0	0	1	0	0	0	0	0	0	1
	Financial awareness	1	0	0	0	0	0	0	0	0	0	1
	Basic information	0	0	0	1	0	1	1	0	0	0	3

Figure X: Table of number of courses matching IRL sub-categories

IRLs sub-categories

Survey Table IRLs *vs* Training

- Indicate whether the training matches each of the IRL's sub-category,

DATA

Sub-category	Value
Pre-Digitisation curation	17
Digitisation	21
Standarisation interoperability	22
Data curation	19
Databasing	26
Data sharing	18
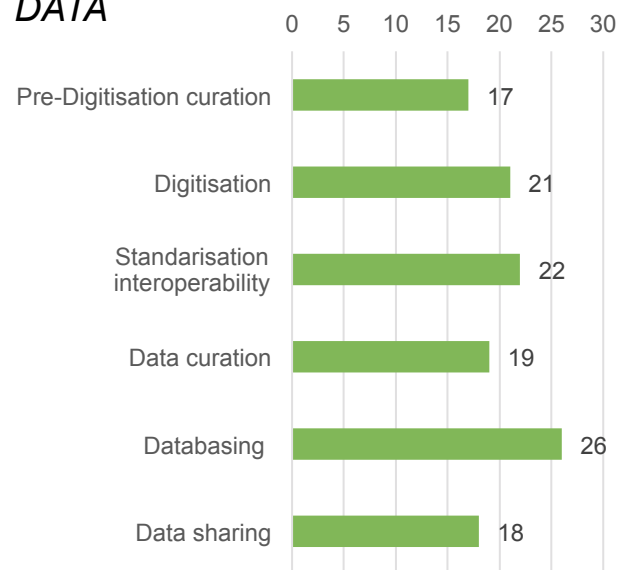
Graphic X : Number of training matching per DATA sub-categorises

TECHNOLOGICAL

Sub-category	Value
Operating	13
Development	6
IT data management	13
e-Monitoring	0
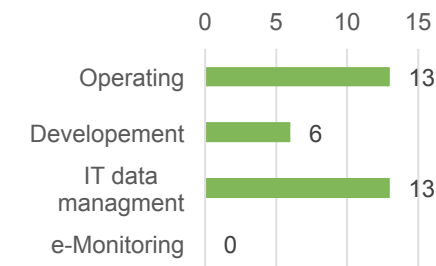
Graphic X : Number of training matching per TECHNOLOGICAL sub-categorises

e-Monitoring: Digital management services (e.g., e-service, helpdesk, mediation services)

First draft of the catalogue

- Title: *Catalogue and recommendations to the development of a proactive, efficient and evolving DiSSCo Training Programme*

TABLE OF CONTENTS

- SECTION 1: SITUATION ANALYSIS OF EXISTING TRAININGS
- SECTION 2: DiSSCo TRAINING NEEDS ASSESSMENT (IRLs)
- SECTION 3: RECOMMENDATION FOR TRAINING DEVELOPMENT & PRIORITIZATION

Proposal Phase 2015 - 2018	Preparatory Phase 2018 - 2021	Construction Phase 2020 - 2023	Transition Phase 2023-2024	Operation 2025 – 2035

The DiSSCo RI works for the **digital unification of all European natural science assets under common curation and access policies and practices** that aim to make the data easily Findable, more Accessible, Interoperable and Reusable (FAIR).

The catalogue will be on the DEST platform
https://cetaf.org/dest/
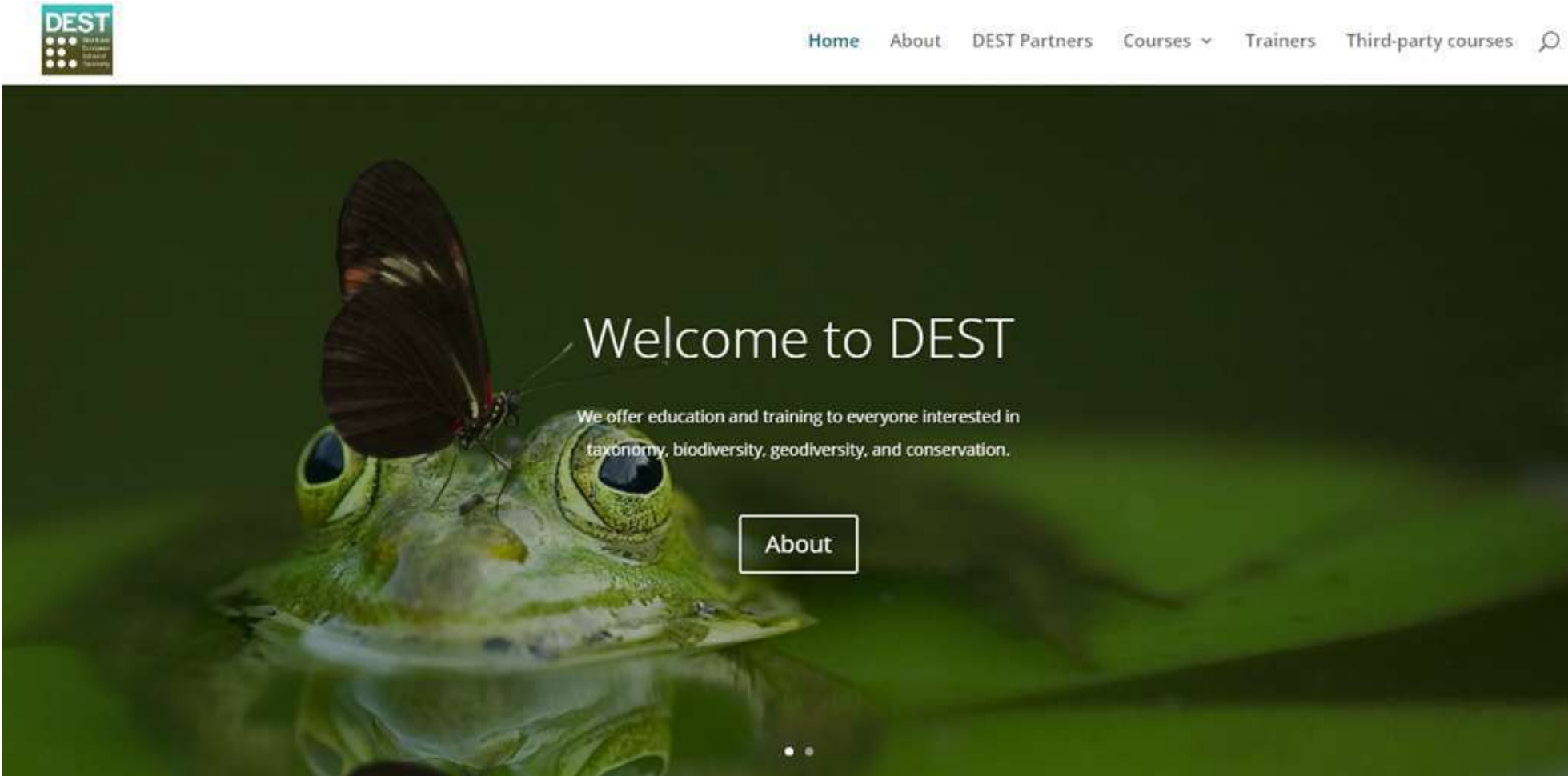
"Bringing the irreplaceable data stored in natural science collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

The preparatory phase project of DiSSCo Research Infrastructure - Distributed System of Scientific Collections

Training Strategy Session / WP2

https://www.dissco.eu/prepare/

Competency and capability – WP3 Task 3.1 and 3.3

DiSSCo All Hands meeting session

13.00 (CET) Tuesday 19th Jan 2021

Task 3.1: Improve digital skills & competencies

- MS 3.1 (Dec 2020) 'consultation and analysis' – report delivered also covers aspects of case studies. Report to be summarised today.

- MS 3.2 (Apr 2021) 'prototype dashboards' – for discussion today about what aspects will be most useful to explore.

- MS 3.3 (May 2021) 'case studies' – will be used to update and fill any gaps from the first report

- MS 3.4 (July 2021) 'digital & data competencies framework' – will be the final tools and recommendations for the deliverable.

DiSSCo
PREPARE

H2020-INFRADEV-2019-2
Grant Agreement No 871088

Improving Digital Capability - Case Studies & Analysis

DiSSCo Prepare WP3 – Milestone 3.1

Helen Hardy (NHM London), Anne Koivunen (Luomus),
Aino Juslen (Luomus), Quentin Groom (MeiseBG), Patricia Mergen
(MeiseBG), Frederik Berger (MfN Berlin), Peter Giere (MfN Berlin),
Rui Figueira (ULisboa), Alexandra Cartaxana (ULIsboa)

NATURAL HISTORY MUSEUM

Areas of exploration (from sub tasks):

1. **Competency frameworks for individuals**, looking at digital and data competencies but also management and leadership of people and projects, and working with collections;

2. **Organisational change to support digital capacity/capability** – including communication, policy, governance, outreach, and organisational structure, and developing towards customised recommendations

3. **Digital capacity data** – e.g. previous surveys. Look at how data collection could be automated and/or data be made more machine readable

DiSSCo needs?

- Leadership of projects and people

- Digital, data & technology skills – to develop infrastructure and services and deliver these

- Specialisms – HR, Finance/procurement etc

All suited to a range of size and scope of teams across members, nodes and central organization – balance of consistency and flexibility.

Integrated RI model
Wide access
Lower costs
Faster
New insights
Optimised
FAIR data

User services

DiSSCO
Distributed System of Scientific Collections

ESFRI RIs
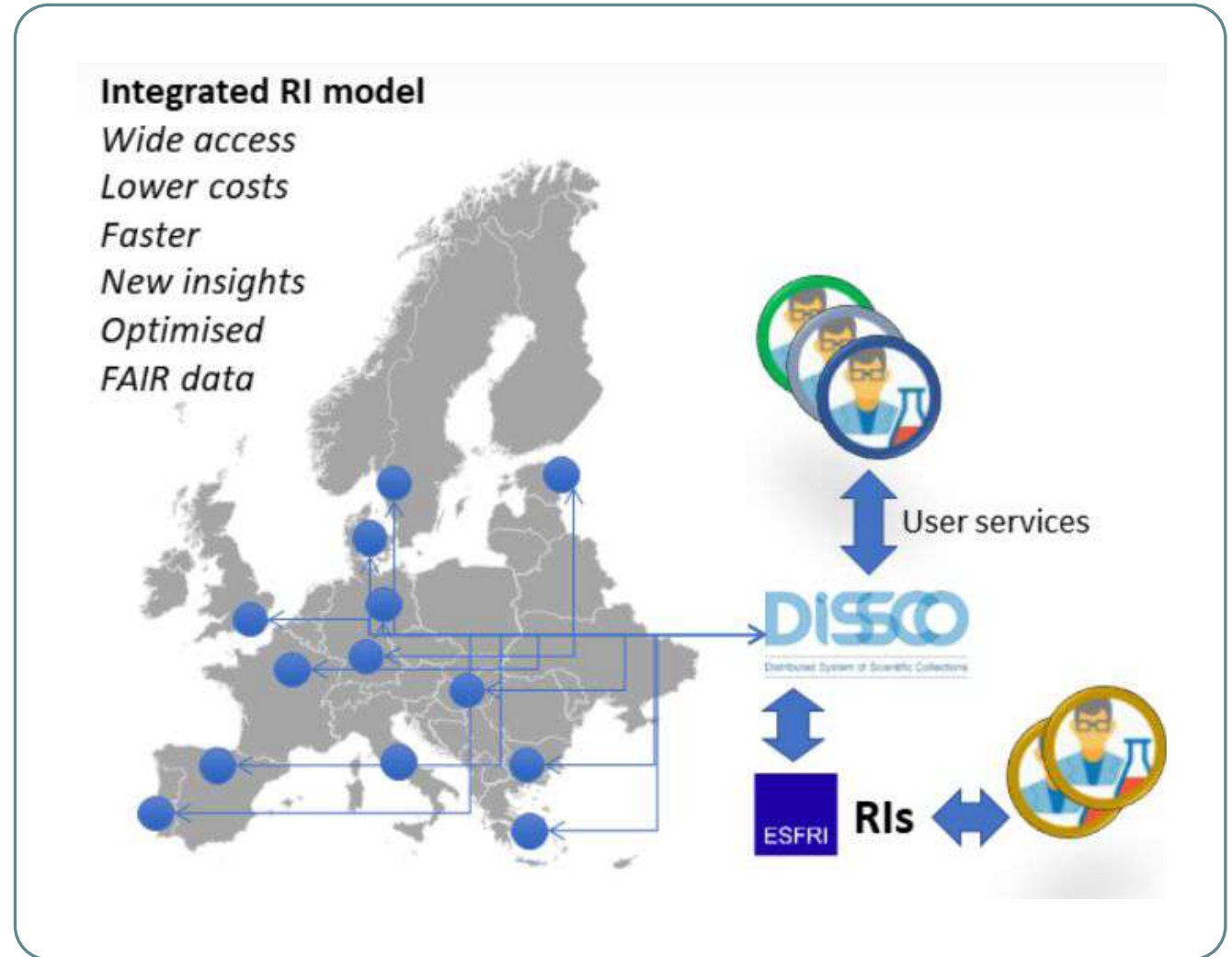
Insights from the prior projects

- No specific previous analysis of competencies or competency frameworks for individuals, or of tools for organisations

- A lot of work on capacity and capability including:
  - Functional and organisational units
  - Existing capacity of organisations
  - Policies (and the categories these fall into)
  - Costs and cost drivers
  - Digitisation workflows and progress
  - Centres of Excellence
  - Training
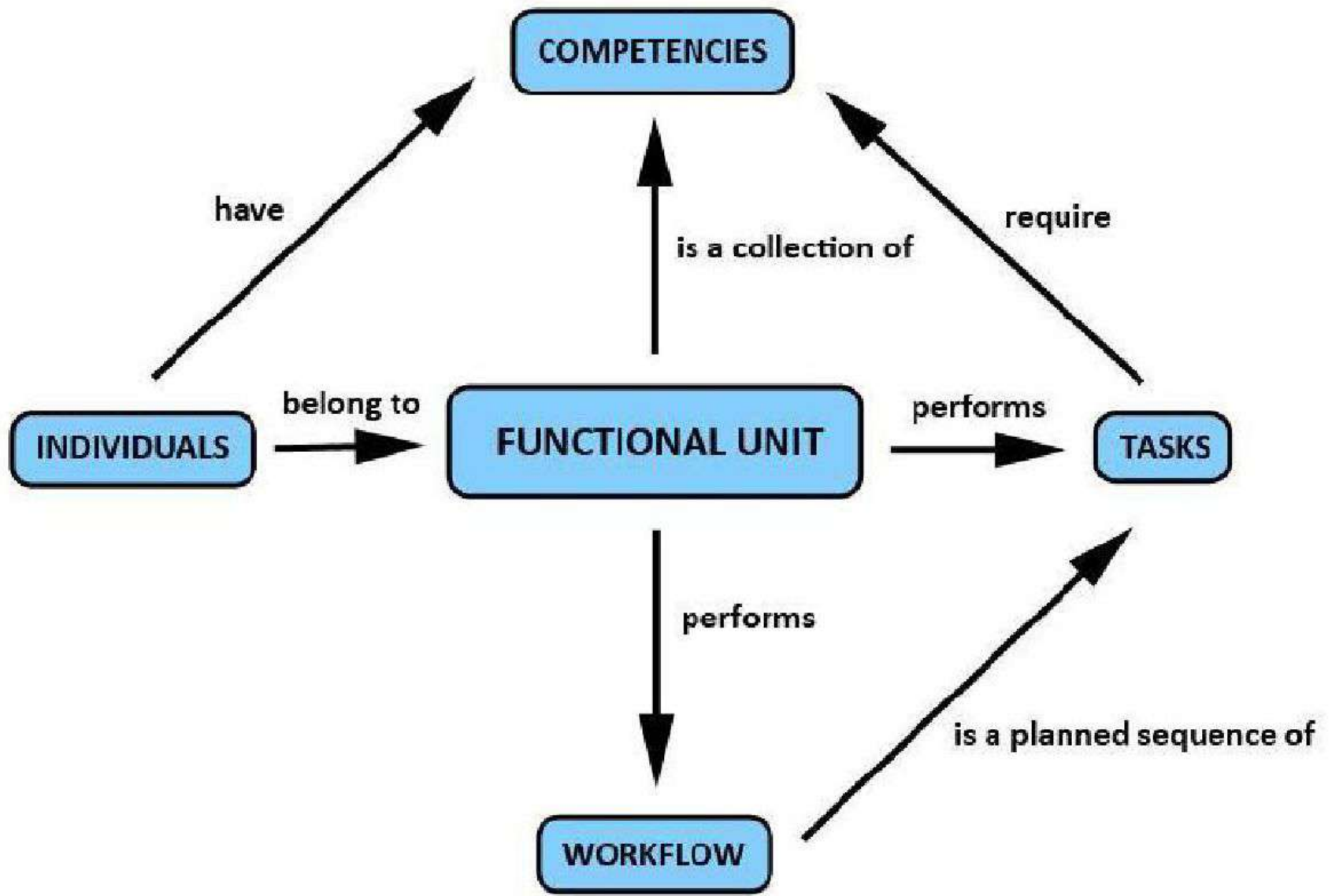
Individual competency frameworks

- EUColComp
- Institutional frameworks
- Open Science frameworks
- EU DigComp
- Data stewardship
- Other sectors e.g. Civil Service
- Skills Framework for the Information Age
- Other examples to explore?

©EU
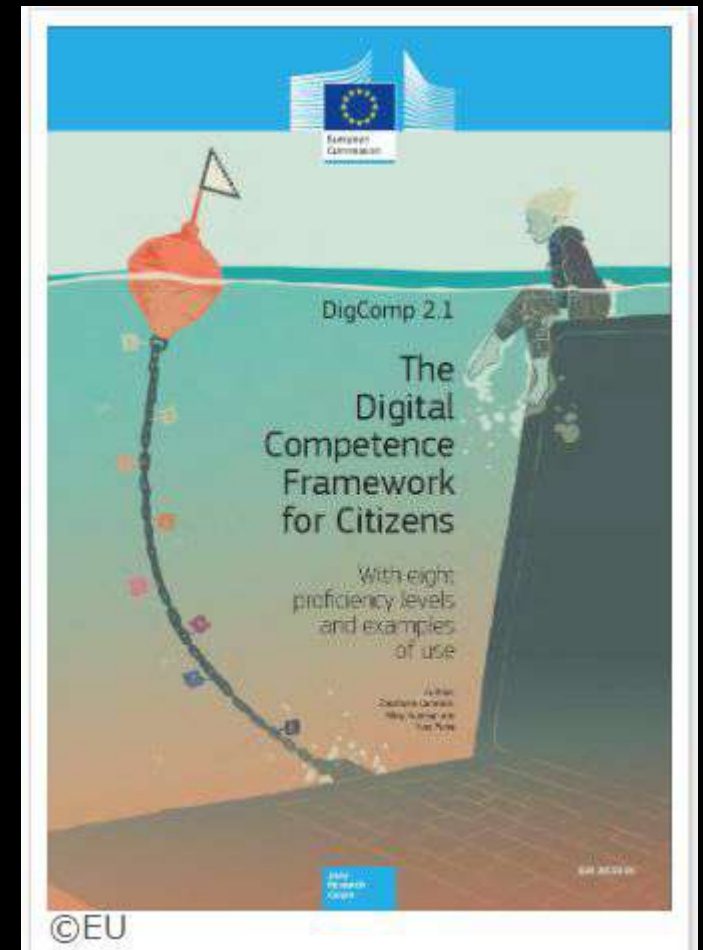
Elements of a competency framework

- The **competencies** themselves - whether limited to behaviours or high level skills such as 'communication', or going wider/deeper into specific areas of knowledge, skills, experiences and qualifications;

- The **levels** at which the competencies are applied, which as a minimum tend to include three levels (like higher/medium/lower or excellent/good/not good), but can also extend to every grade of an organisation (7-8 levels is quite common); and

- The **generic attributes or expectations** which define those levels i.e. how they increase - typically including increasing autonomy, complexity and cognitive burden with increasing seniority.

All part of an ecosystem with role profiles / job descriptions; professions or specialisms and their qualifications; individual objectives; performance management and appraisal policy; recruitment law & practice; promotion / progression, and training / development.

What are the elements of the Success Profile?

Experience

the knowledge or mastery of an activity or subject gained through involvement in or exposure to it.

Strengths

the things we do regularly, do well and that motivate us.

Ability

the aptitude or potential to perform to the required standard.

Technical

the demonstration of specific professional skills, knowledge or qualifications.

Behaviours

the actions and activities that people do which result in effective performance in a job.

Success Profiles

Experience

Ability

Technical

Behaviours

Strengths

Organisational digital maturity/capability

- Institutional approaches and case studies

- External examples e.g. UK Digital Culture Compass

- GBIF

Tracker

This self-assessment tool provides a practical framework for improving your organisation's digital strategy, planning and activities. Once registered, you can assess your current status, set 12-month targets and share reports with colleagues.

USE TRACKER

1. Initial
2. Managed
3. Integrated
4. Optimising
5. Transforming
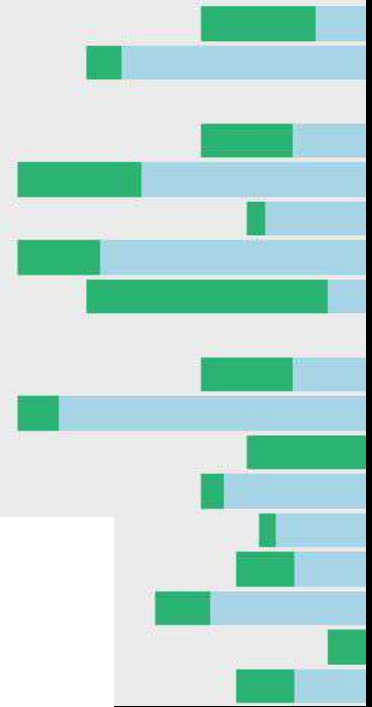
Organisational Change in Finnish Museum of Natural History, Luomus - Case Study

1. **Strong support from the open data policy implemented in the EU and especially on the national level.**

2. **Clear vision from management.**

3. Strong staff engagement in the process of creating the guiding documents endorsing open science e.g. in policies and implementation plans.

4. **Solid process of setting concrete goals, with line of sight from university's strategy → institution level → unit level → team level → personal level.**

5. Reshuffling the pre-existing resources in new ways, with a new unit focusing on different aspects of digitalisation and new teams specialising in digitisation, ICT and species information.

6. External funding for some additional personnel, investments in digitisation equipment, the establishment of the Finnish Biodiversity Info Facility (FinBIF - https://laji.fi/en) as a repository for FAIR data dissemination.

7. An effective collection management system, **Kotka**, that is well suited to our institution´s needs.

8. Suitable training for the staff.

Pros and Cons

The Concrete Goals

The Vision

The Supportive Operating Environment
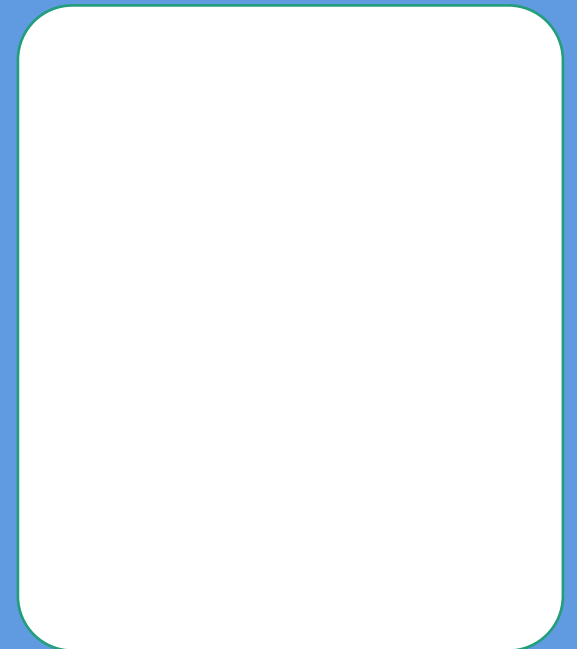
Data, and finding skills/capacity

- How can we find organisations or colleagues who have the skills and capabilities that DiSSCo needs?

- And/or how can we standardise vocabulary for frameworks, tools and data entry?

- Various sources but risks around whether these are up to date and sustainable.

Key discussion points

1. What competency frameworks or digital capability tools are used in your organisation?

2. Are there other examples we should look at?

3. What should be the priorities for this task – what will be most useful in practice and why?
   - Help to navigate existing frameworks and tools?
   - A new competency framework for DiSSCo? High level or detailed?
   - Generic role profiles for key roles?
   - A DiSSCo-specific digital maturity self-assessment tool for organisations?
   - Work using e.g. LinkedIn and ORCID APIs to assess data viability?
   - Work to establish key vocabularies?
   - Other?

Task 3.3 (from sub tasks)

- Develop a formal **staff secondment procedure** that balances the needs of the donor and host institutions across different sizes and types of organisations. Example models to explore include the creation of a central secondment fund across DiSSCo, routes for the host institute to pay, and route for the donor institute to pay – and run a **pilot** involving part of the DiSSCo technical architecture;

- Develop cross- institutional work practices and tools to support teams of individual staff working collectively on common tasks - examining best practices from other research communities, different project management approaches, and the communication infrastructure necessary to deliver effective **distributed team working**.

Task 3.2 Collate, refine and implement best practices for data mobilisation at the institutional level to develop the DiSSCo plan for data mobilisation and curation pipelines

Monday, January 18th 2021
Speaker: Laurence Livermore (NHMUK)
Partners: Luomus, MeiseBG, MfN, RBGE, ULISBOA
Note taking: Helen Hardy (NHMUK)

Funded by the Horizon 2020
Framework Programme
of the European Union

January 18th 2021

Task Overview

How do you best prepare collections for digitisation, digitise them, curate the associated data, publish this information and measure the outputs? What are the options and rationale for different types and sizes of collections, when should this be outsourced and what different project management approaches are most appropriate in this range of circumstances?

This task seeks to address these questions, describing and refining best practices and building on a substantial investment from prior and current projects and feeding these into DiSSCo Prepare WP8. Consolidating what is known into a community-edited manual (supported by WP5), and other relevant platforms, WP3 will streamline the reuse and implementation of these procedures and enhance digitisation capacity across the DiSSCo collection-holding organisations.

Subtasks

3.2.1 Standard Operating Procedures for digitisation (SOPs) [MS3.5]

3.2.2 Standardised Extract Transform and Load (ETL) procedures [MS3.6]

3.2.3 Pre-Digitisation Curation [MS3.7]

3.2.4 Digitisation Monitoring [MS3.8]
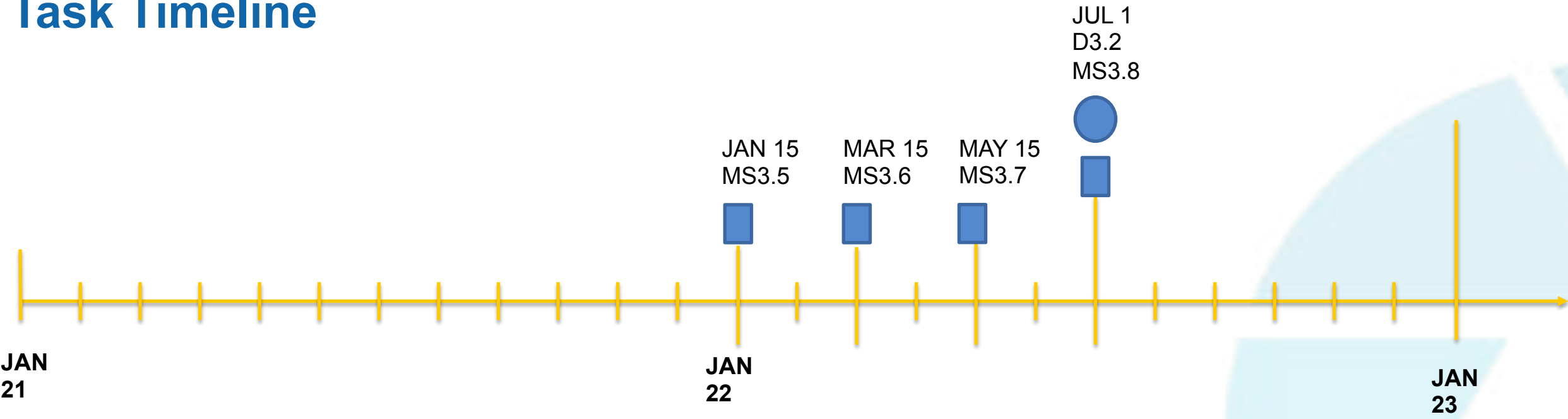
Discussion Points

1. Current state of institutional digitisation

2. General state of natural science collections digitisation

3. Task logistics

See document links:

Planning https://docs.google.com/document/d/12ywJGyo2_4ps8ogBwJ0HP-arVlWld-BpZ8tmZMd1LvA

Notes https://docs.google.com/document/d/1EWyEHNNdm6UoL0gnffmCH2pliTE0HH0kJteJeet6C1M

Funded by the Horizon 2020
Framework Programme
of the European Union

AHM – WP4.1 – METHODOLOGY ASSESSMENT
Thursday, January 21st 2021

Convener: Michel GUIRAUD (MNHN)

January 21st 2021

1

MEETING AGENDA

TIME	TOPIC
9:00 – 9:10	Presentation of WP4.1 (DiSSCo Cost Book)
9:10 – 9:30	Introduction to indirect costs
9:30 – 10:00	Discussion on indirect costs
10:00 – 10:30	Methodology for the calculation of DiSSCo's indirect costs

HOW THE COST BOOK WORKS?
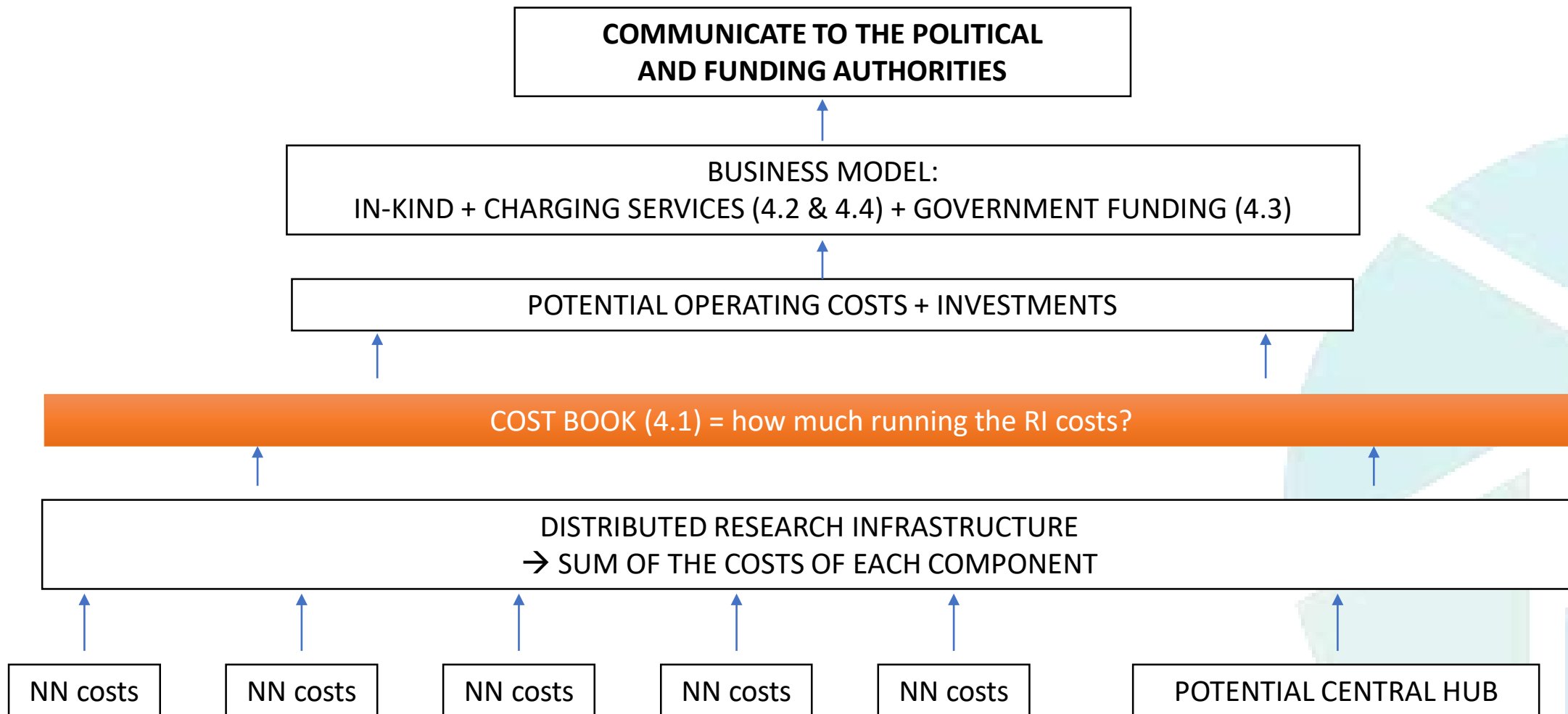
DIFFERENCE BETWEEN DIRECT COSTS AND INDIRECT COSTS

DIRECT

STAFF (scientific/technical/administrative essential to the operation of RI)

EQUIPMENT directly used by the RI

MAINTAINING this equipment in operational conditions

CONSUMABLES

STUDIES AND SERVICES

INDIRECT

GENERAL ADMINISTRATION (management, training, communication, legal services, finance department, HR, etc.)

GENERAL COMMON RESOURCES (catering and accommodation, postal services, handling, public transport)

BUILDING INFRASTRUCTURE (electrical installations, heating, water and air installation, cleaning, etc.)

SECURITY / SAFETY (health and safety, security, etc.),

IT OPERATION AND MAINTENANCE

WHAT ARE INDIRECT COSTS?
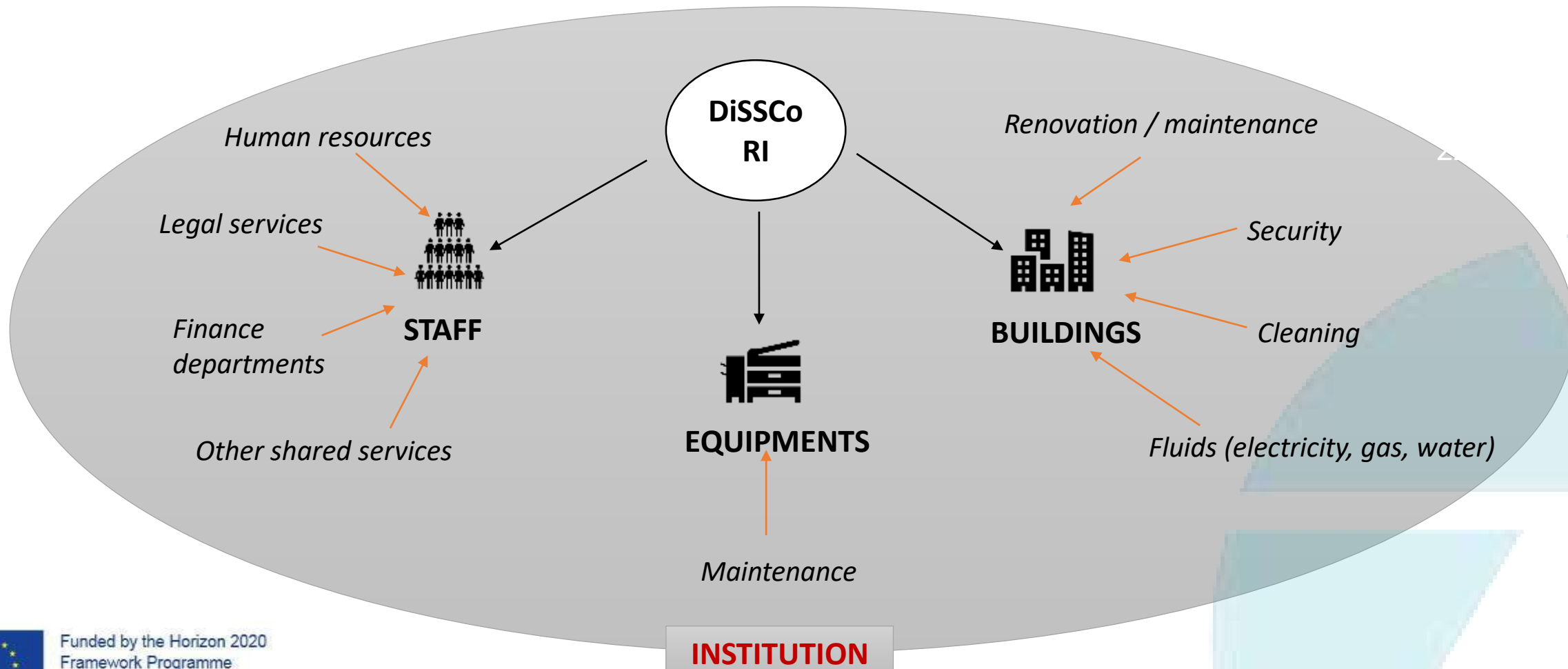
HOW CAN WE ESTIMATE THOSE INDIRECT COSTS?

ASSESS THESE COSTS WITHIN THE INSTITUTION
How much HR services cost per year? (etc.)

DEVELOP A RATIO: COST PER STAFF MEMBER OR PER SQUARE METRE

= GROSS ANNUAL SALARY CHARGED OF THE STAFF INVOLVED IN THE PROJECT * RATIO
or / and
= SQUARE METRES OCCUPIED BY THE PROJECT * RATIO

FIRST RESULTS (2019)

	NHM	ISBAS	MNHN	UniFI	BGBM
ADMINISTRATION	9 796 640,00 €	162 360,00 €	15 598 155,00 €	No info	639 395,00 €
BUILDING	14 148 958,88 €	148 600,00 €	7 654 703,00 €	No info	3 787 062,00 €
STAFF	936	136	2685	45	213
M2 (without gardens)	127 601 m2	4 912 m2	129 000 m2	13 800 m2	12 477 m2
SUPPORT COST / PERSON / YEAR	10 466 €	1 193 €	5 809 €	No info	3 001 €
SUPPORT COST / M2 /YEAR	110,88 €	30,25 €	59,34 €	No info	303,51 €

DISCUSSION

- Are you familiar with a system for calculating indirect costs specific to your institution?

- If so, how was it developed?

- Does it take into account all these costs?

- If you do not have a personalized system, do the overheads offered by funders balance the indirect costs of your institutions?

METHODOLOGY
WHICH METHOD IS ACCESSIBLE TO YOUR INSTITUTION?

a) In-house calculation of indirect costs

b) MNHN proposes a ratio to be applied to direct costs

c) Development of a differentiated ratio according to the type of institution

d) Other proposal

"Bringing the irreplaceable data stored in natural science

collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

The preparatory phase project of DiSSCo
Research Infrastructure - Distributed
System of Scientific Collections

https://www.dissco.eu/prepare/

Task 4.2 Cost model for charging services

Speaker: Laurence Livermore (NHMUK)
Partners: IRSNB, MeiseBG, MNHN, +Technical Subcontractor(TBC)
Note taking: Lizzy Devenish (NHMUK)

January 21° 2021

1

Session Aims

- Task overview

- Discuss task scope & relationship with other tasks/projects

- Discuss subtasks & subtask leadership

- Agree next steps

"Task 4.2 will identify and fully cost the range of services to be provided by the DiSSCo RI"

Broad scope including:
- Data mobilisation, processing activities, access and exploitation
- Common services core to DiSSCo RI operations
- Platform to store cost data & manage access
- Identify users and market niches to verify service offers meet real demands

Task timeline
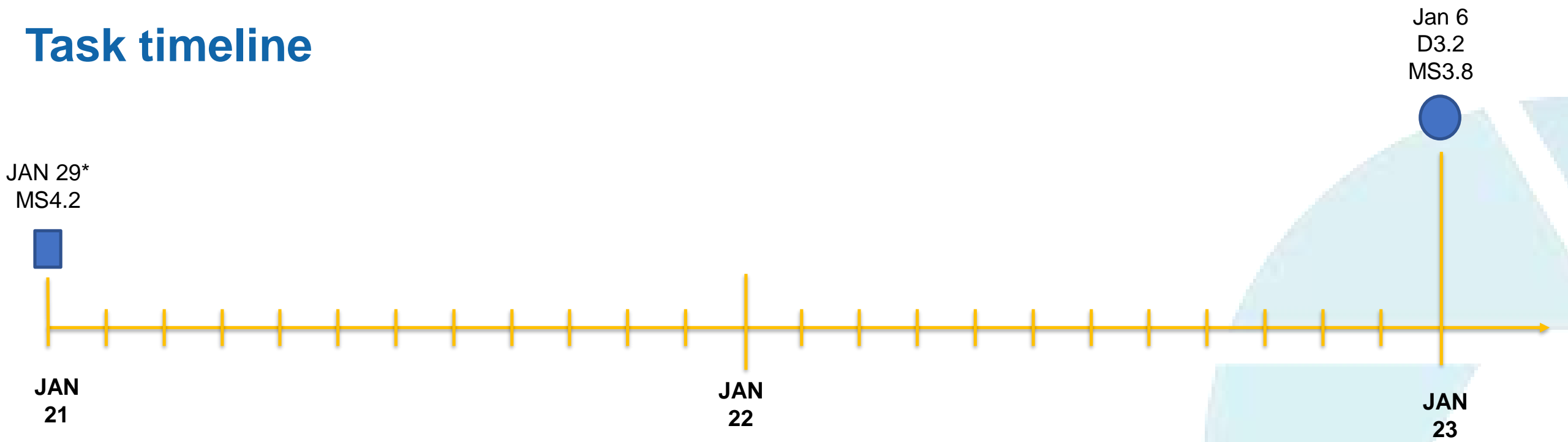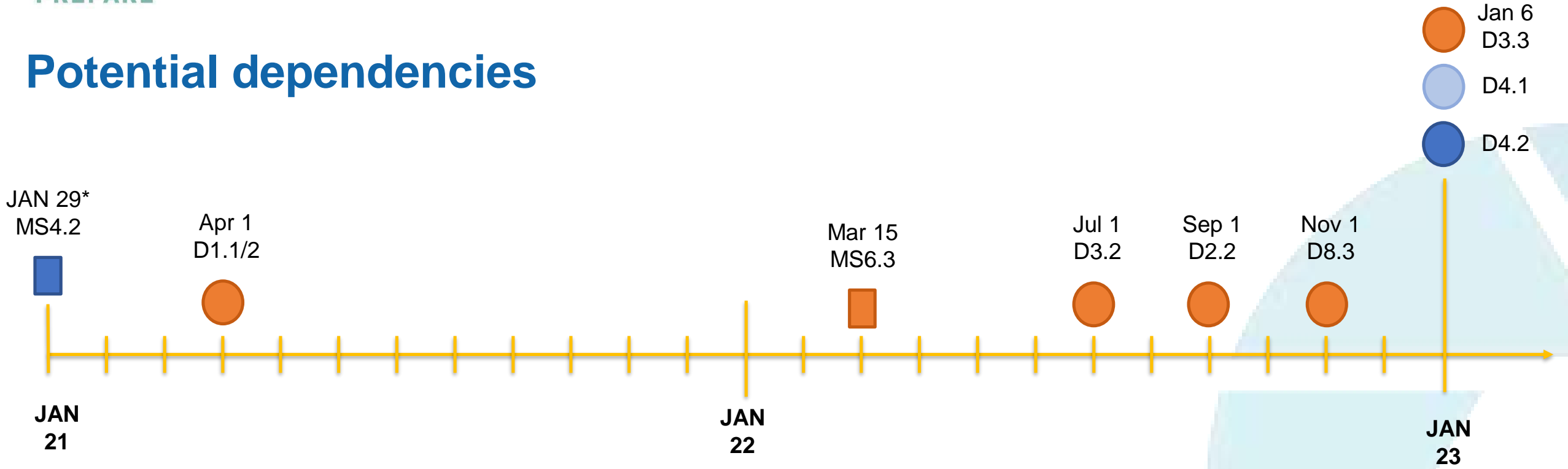
WP3 Task 3.3 includes:

Develop a formal **staff secondment procedure** that balances the needs of the donor and host institutions across different sizes and types of organisations. Example models to explore include:

- Creation of a **central secondment fund** across DiSSCo (likely to be needed to make this work);
- Routes for the host institute to pay; and
- Route for the donor institute to pay.

Run a **pilot** involving part of the DiSSCo technical architecture.

Task 3.1 is also looking at competencies and capability – may help to identify circumstances where outsourcing is relevant.

Proposed subtasks

- 4.2.1 Digitisation costs
- 4.2.2 Common DiSSCo Services
- 4.2.3 Cost model platform*
- 4.2.4 User needs & testing*

*Denotes involvement of technical subcontractor

Next steps

- Delimit scope of Task 4.1 and Task 4.2

- Read and discuss recommendations from ICEDIG reports

- Review links and dependencies with other WPs

- Discuss Cost Model platform with DiSSCoTechgroup

- Any other points to consider?

Session Task 5.1

DiSSCo Knowledgebase

Julia Pim Reis, Sabine von Mering, Mareike Petersen & Falko Glöckler

Museum für Naturkunde

Central and freely accessible DiSSCo Knowledgebase

Store all research outputs from DiSSCo-linked projects in one place but also act as a reference for further building blocks relevant for the DiSSCo Research Infrastructure (RI).
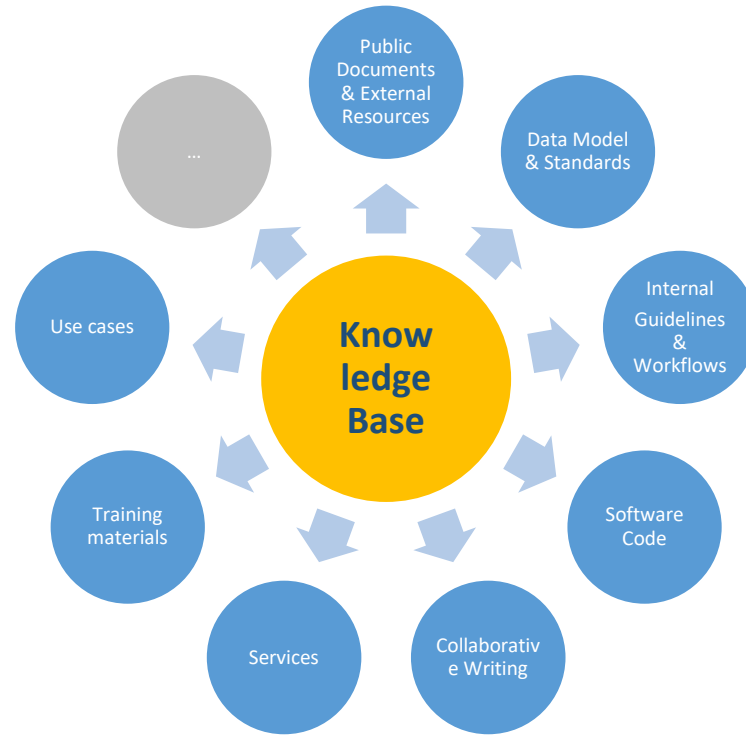
Figure 1: Information Types in the DiSSCo Knowledgebase. Expected cluster of information categories (blue dots) based on DPP Project outcomes and relevant external resources. The format of resources varies within and among information types.

Process of System Decision
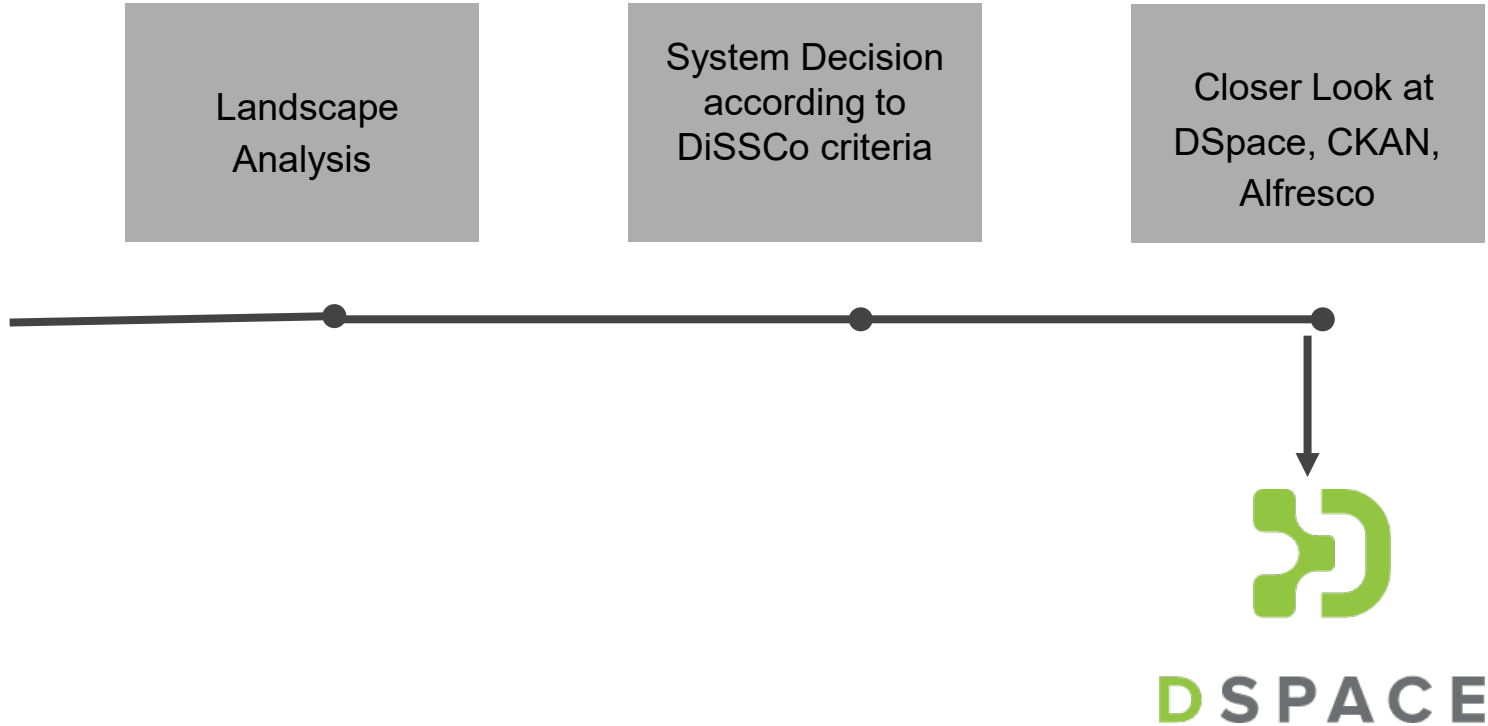
What is DSpace?

DSpace is an **open source** software package typically used for creating open access **repositories** for scholarly and/or published digital content. While DSpace shares some feature overlap with content management systems and document management systems, the DSpace repository software serves a specific need as a digital archives system, focused on the **long-term storage**, **access and preservation** of digital content.

Large user community: >1000 instances of DSpace running worldwide

Blogpost in DiSSCo Tech blog

The DiSSCo Knowledgebase

Authors: Mareike Petersen*, Julia Pim Reis*, Sabine von Mering*, Falko Glöckler*

* Museum für Naturkunde Berlin, Germany

Introduction

As an initiative formed by public research institutions, DiSSCo is committed to Open Science. We believe that Open Science not only makes the scientific work more transparent and accessible but also enables a whole new set of collaborative and IT-based scientific methods. Therefore, the outputs of our common research projects are openly available as much as possible and research data easily Findable, more

https://dissco.tech/2020/12/18/the-dissco-knowledgebase/

First version of DiSSCo Knowledgebase

Demonstration Time!
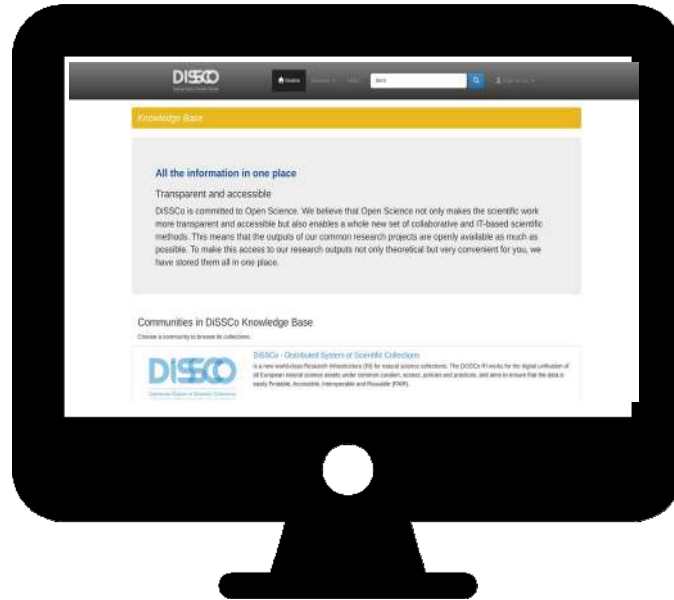
DiSSCo PREPARE
Virtual meeting

Next steps

- improve style (corporate design)

- implement session feedback regarding user interface

- (productive) launch of DiSSCo KB: mid February 2021

- implement open ID mechanisms (ORCiD or other)

- automatic DOI assignment: second half of 2021

All Hands Meeting 1
5.4 - DiSSCo update

21 Jan 2021

Wouter Addink- Naturalis

DiSSCo Prepare WP5.4

Modernising technical infrastructure for

science data mobilisation and publication

Funded by the Horizon 2020
Framework Programme
of the European Union

AHM1 / T5.4

The three building blocks of the DiSSCo technical infrastructure

For more information: https://dissco.tech/

Repositories with data provided by the DiSSCo participating institutions
- Trusted repositories: Institutional repositories and global thematic repositories such as GBIF, GeoCase.
- Connections to third-party repositories like genetic sequence and literature databases.
- All data that can be linked to collection objects (specimens) are in scope.

Digital Object Infrastructure
- FAIR Digital Objects (DO) architecture (with Digital Specimen as core objects)
- DOIPv2 (Digital Object Interface protocol) and Handle protocol (with "powered by DOI" PIDs for the Digital Specimen)
- CORDRA software for the Natural Sciences Identifier Registry (nsidr.org)
- AAI infrastructure compatible with European Open Science Cloud (EOSC)
    -> CORDRA + KeyCloak pilot implementation

Community Services
- DiSSCo e-Services to discover, consume and interact with the federated data
- Services provided in collaboration with other research infrastructures

Loans and Visits System (ELViS)

a one-stop shop for access to the collections, providing both physical access and virtual access by digitisation on demand

Collections Monitoring Dashboards

Dashboards showing the digitisation status and usage of the collections

Knowledge base

- digitisation resources
- protocols
- manuals
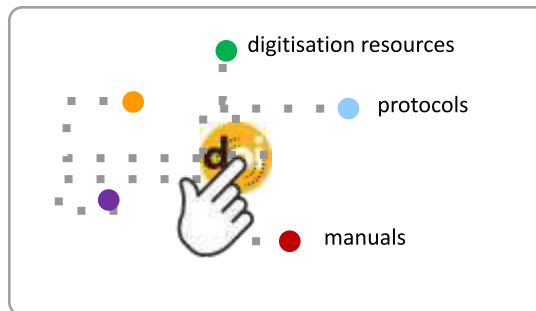
access protocols, digitisation resources, manuals and other documents as FAIR digital objects for direct integration with the other services

European Curation and Annotation System (ECAS)

- Genomic data
- Biochemical data
- Morphological data
- Geographical data
- Taxonomic Information
- Species Interactions data
- Ecological data
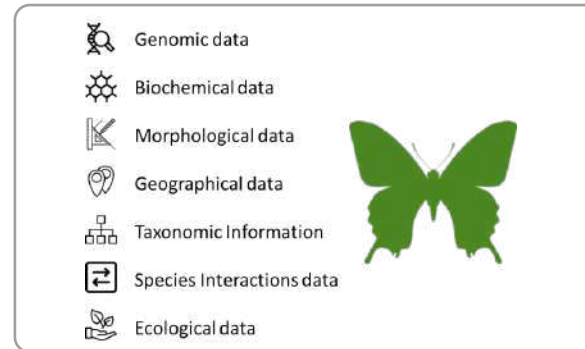
community curation of the digital specimens, the digital twins of the physical specimens that contain all digital information derived from the specimens and bring the specimens in connection with species, habitat and other related data

Specimen Data Refinery (SDR)

Segment and crop parts of images

Group similar specimens and labels (based on size, shape, colour, landmarks)

Measure specimens and labels
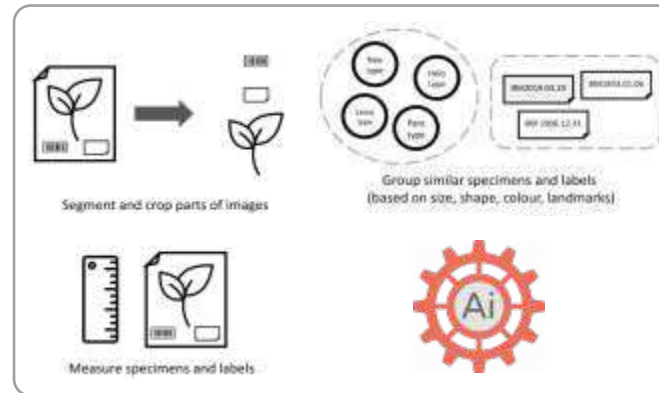
digitisation services to extract, enhance and annotate data from specimens digital images

DiSSCo PREPARE

Community services (core e-Services)

The digitisation status of collections in Europe today
(CETAF & DiSSCo countries)

only about **4.8%** of the estimated 1.5 Billion specimen in collections in Europe are digitised and published in GBIF as preservedSpecimen, fossilSpecimen or livingSpecimen
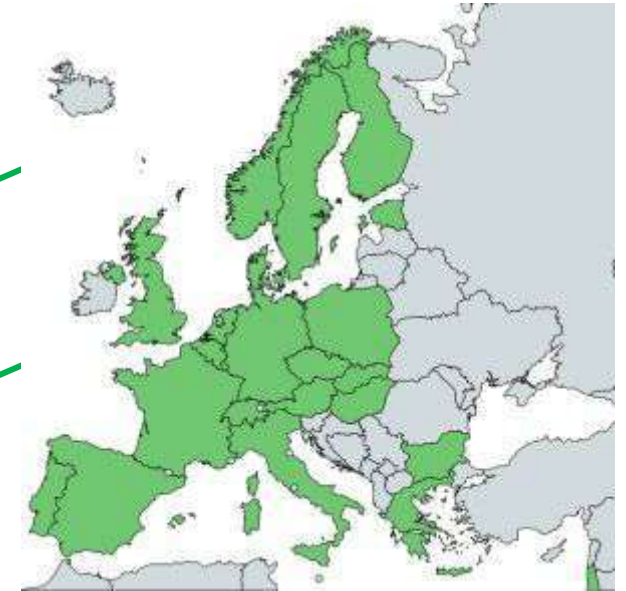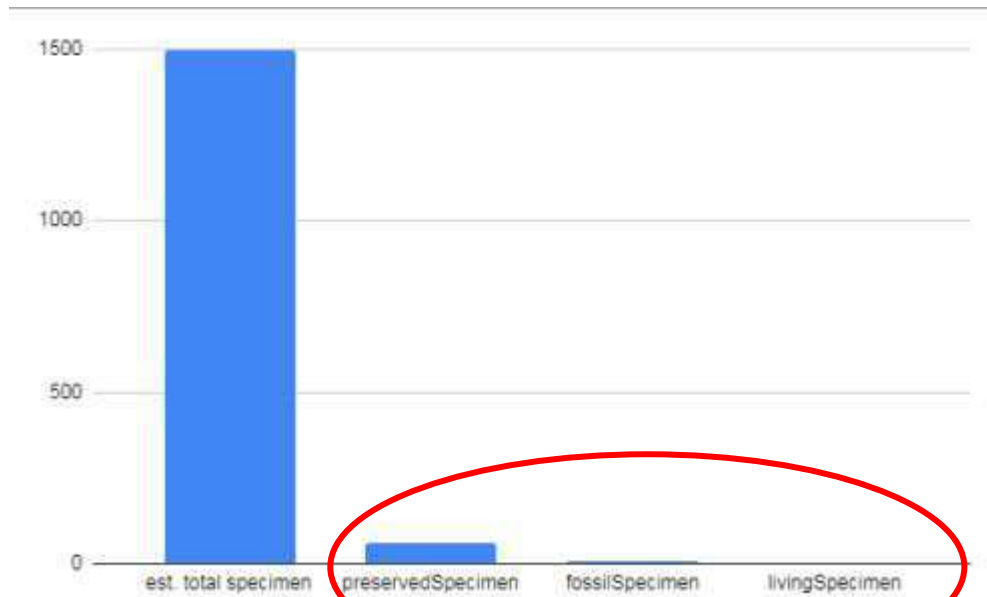
Digitised and published in GBIF, BioCase

Collection Digitisation Dashboard pilot

https://rebrand.ly/synth-cdd

ELViS european loans and visit system

- Integrated system for Europe to support loans, visits, applications for digitisation on demand, and to track outputs.
- DiSSCo e-Service for access to European collections
- A future service in the European Open Science Cloud (EOSC)
- Developed in SYNTHESYS+ project, EC funded

-> Used for SYNTHESYS+ Virtual Access (Digitisation on Demand) and future Transnational Access calls

Includes:
- Institution description based on GRID.AC and CETAF Register
- Facilities & instruments
- Collection descriptions (planned)
- Specialists descriptions (planned)

Known to contain types ● Yes ○ No
Holds specimen ● Yes ○ No

Contact person

Visiting address

Street and number Street Number
Zip code and City Zip code City

Preservation

Discipline Botany

☑ Pressed and dried ☐ Fluid preserved ☐ Cryopreserved/frozen 80C
☐ Dried ☐ Microscopic slides ☐ Spore print

Discipline Zoology vertebrates

☐ Dried and pinned ☑ Dried - assembled ☑ Dried - not assembled
☑ Fluid preserved ☑ Microscope slides ☐ Cryopreserved / frozen -80C
☑ Dried assemblage ☑ Dried - not assemblage ☑ Fluid preserved
☑ Microscope slides ☐ Cryopreserved/frozen - 80C

Discipline Zoology invertebrates

☑ Dried and pinned ☐ Dried - assembled ☑ Dried - not assembled
☑ Fluid preserved ☑ Microscope slides ☐ Cryopreserved / frozen -80C
☐ Dried assemblage ☑ Dried - not assemblage ☑ Fluid preserved

Global Workflow

Knowledgebase

To be launched soon at
http://know.dissco.eu

DSPACE implementation with full text search, REST API, custom metadata, versioned documents, DOIs.

To store all research outputs from DiSSCo-linked projects in one place (hierarchical structure)

structured technical documentation of identified DiSSCo technical building blocks, (e.g. web services, PID systems, controlled vocabularies, ontologies, data standards)

Focus on human readability but metadata will be machine-readable.

See also: https://dissco.tech/2020/12/18/the-dissco-knowledgebase/

Specimen Data Refinery
Piloted in SYNTHESYS+

Pipelines and services to extract, enhance and annotate
data from specimens digital images

Segment and crop parts of images

Group similar specimens and labels
(based on size, shape, colour, landmarks)

Measure specimens and labels
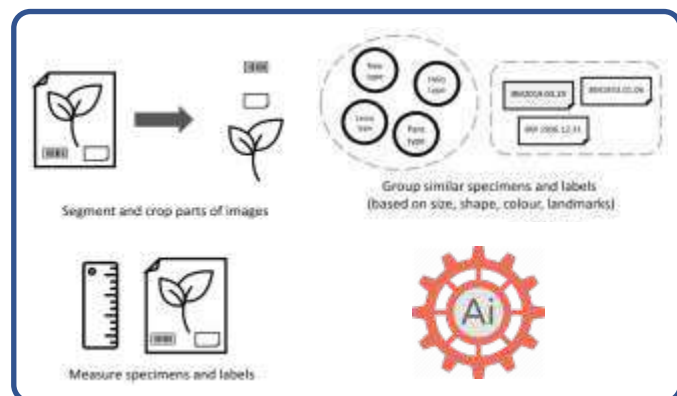
Workflow and data infrastructure integration

- Galaxy (from ELIXIR) and WorkflowHub (from ELIXIR)
  - Both implement RO-Crate
- Digital Specimen (openDS) repositories
  https://github.com/DiSSCo/openDS#repository-structure
- OpenAIRE integration
  - Galaxy will be an openAIRE project, to put their outputs into Zenodo using RO-Crate
  - Push workflows from WorkflowHub into openAIRE through the EOSC-Enhance project

  - RO-Crate support for Zenodo through the CS3MESH4EOSC project.

BICIKL - Biodiversity Community Integrated Knowledge Library
--> a new project (36 Months) to start in May 2021

- BICIKL will initiate and build a new European starting community of key research infrastructures in the domain of biodiversity
- Provision of access to data, associated tools and services at (1) each separate stage of, and (2) along the entire research cycle.
- New methods and workflows for an integrated access to harvesting, liberating, linking, accessing and re-using of sub-article-level data (specimens, material citations, samples, sequences, taxonomic names, taxonomic treatments, figures, tables) extracted from literature.
- Tools for seamless linking and usage tracking of data along the line: specimens → sequences → species → analytics → publications → biodiversity knowledge graph → re-use.

Will deliver key components that support the FAIR Digital Object Architecture of DiSSCo:
- a pan-European PID system for Digital Specimens
- functional links between specimens and other data classes
- link and validate entities discovered in specimens through wikidata
- Specimen information provided as FAIR Digital Objects

Timeline technical developments

Identification Gaps
and User stories,
Initial infrastructure Design
Demonstrator

Standards and specifications development,
Pilots,
Refining Infrastructure Design

Start construction

| 2020 | 2021 | 2022 | 2023 |

Timeline technical developments (2)

DiSSC PREPARE

| 2020 | 2021 | 2022 |

ELViS MVP

ELViS V1 (TA+VA)

ELViS V2 (+ Loans, IAA, pilots)

WP1: Life Science and Earth Science user stories

T6.3: technical interface requirements end–user services

T5.3: semantic enhancement and interoperability

T5.4: modernizing technical infrastructure for data mobilisation

Standards development (MIDS, CD, Open DS)

Digital Specimen Architecture Pilot
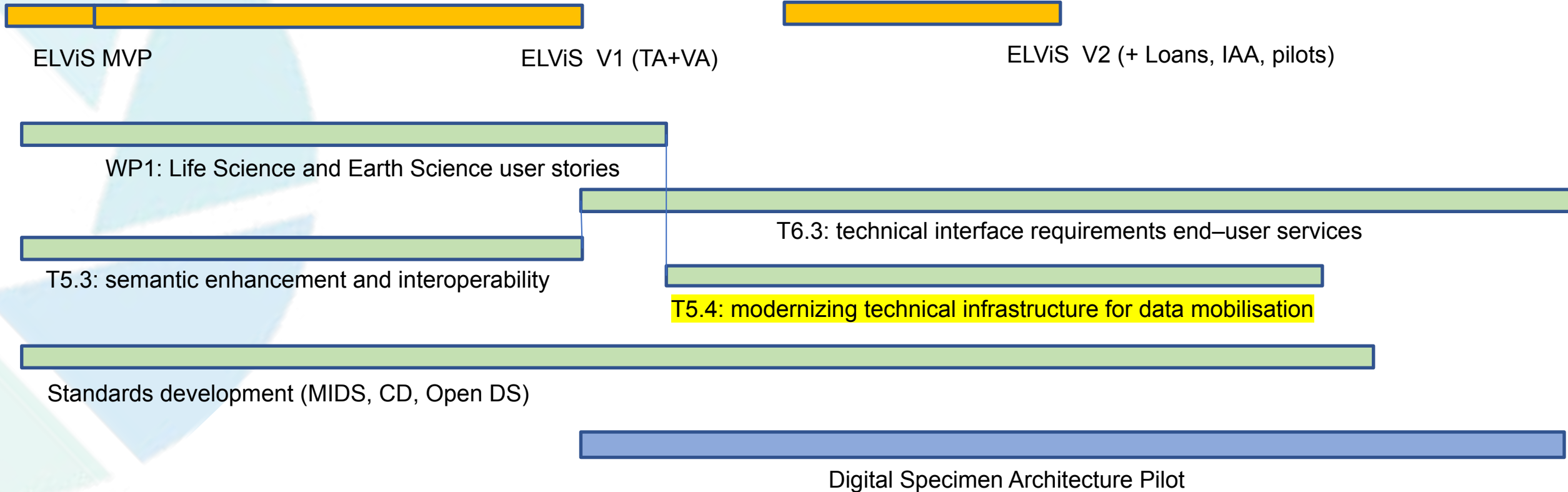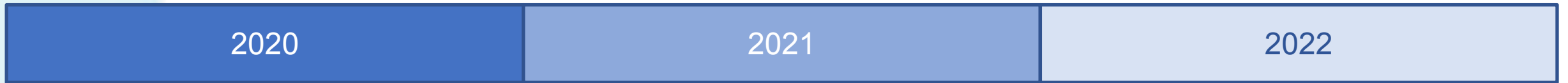
DiSSCo Digital Specimen Architecture pilot (2021-2022)

Objective:

- To further develop and transition the nsidr.org demonstrator into a Digital Specimen pilot to showcase and test the novel principles of a FAIR Digital Object infrastructure (FAIR DO)

Requirements:

- Need to support the OpenDS specification for Digital Specimens, currently being created in the DiSSCo Prepare project.
- It should also get a real dataset, scaled up from a few specimens to millions of specimens. to get experience with operation at scale but also to make it interesting for early adaptors.
- The pilot will need to support stories that showcase advantages the new infrastructure will bring.
- Further components need to be developed when decisions are made on a PID scheme for the objects, to support early implementation in DiSSCo e-services already in development.

Dependent on DiSSCo Prepare T6.2 and 5.2

Stories & demonstrators development

3 demonstrators/pilots:
- CMS integration demonstrator –related to DPP T6.1
- Specimen Data Refinery demonstrator – related to Synthesys+ T8.3, T8.4
- Collection Digitisation Dashboard automation & data management – related to DPP 6.1

Stories procurement:
- 2-3 Stories to be developed by DiSSCo nodes in the form of video's, proof of concepts, interactive Jupyter notebooks
- Related to DiSSCo Prepare WP6.3
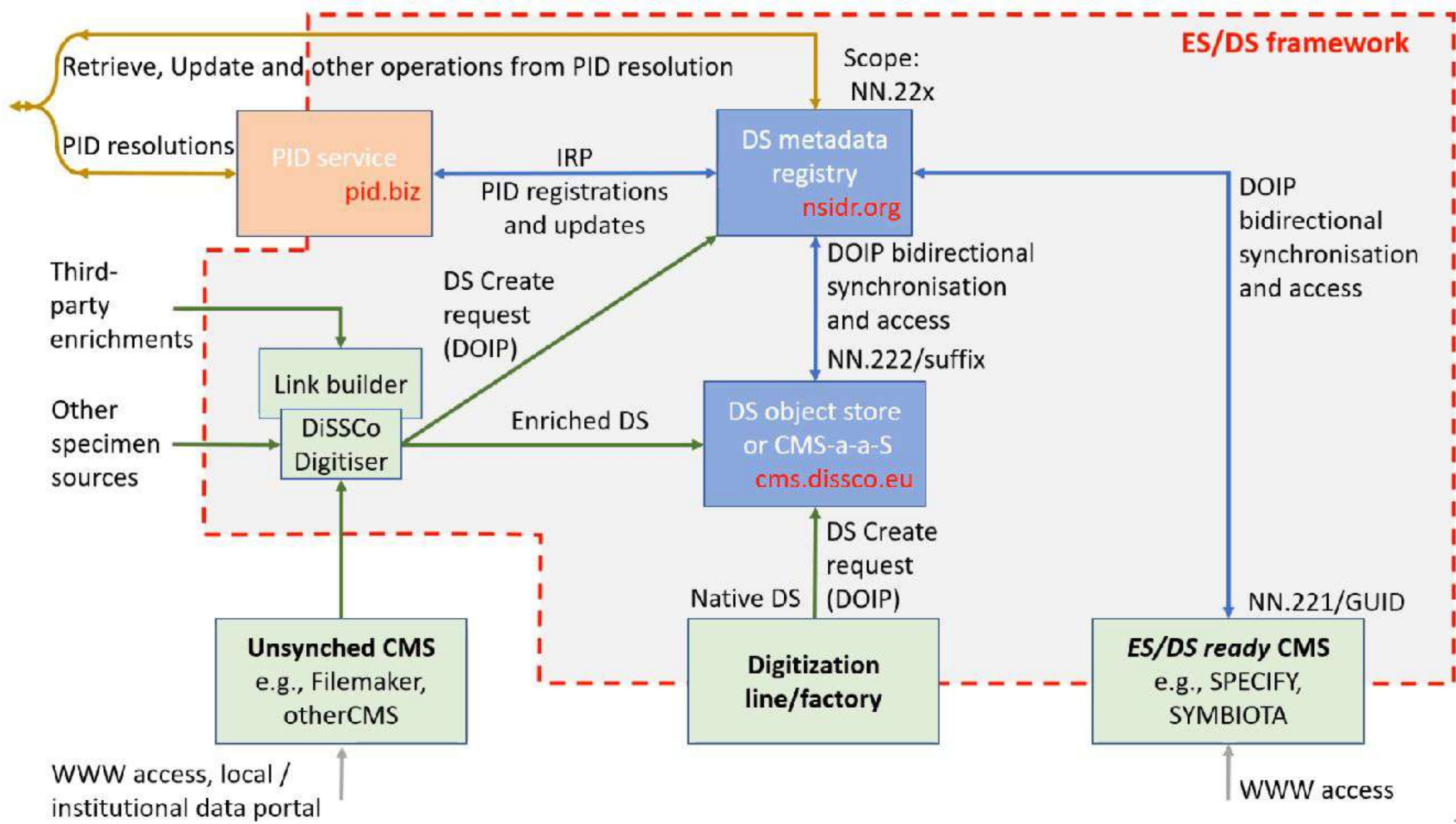
DS classes as specializations of abstract classes, with extending traits
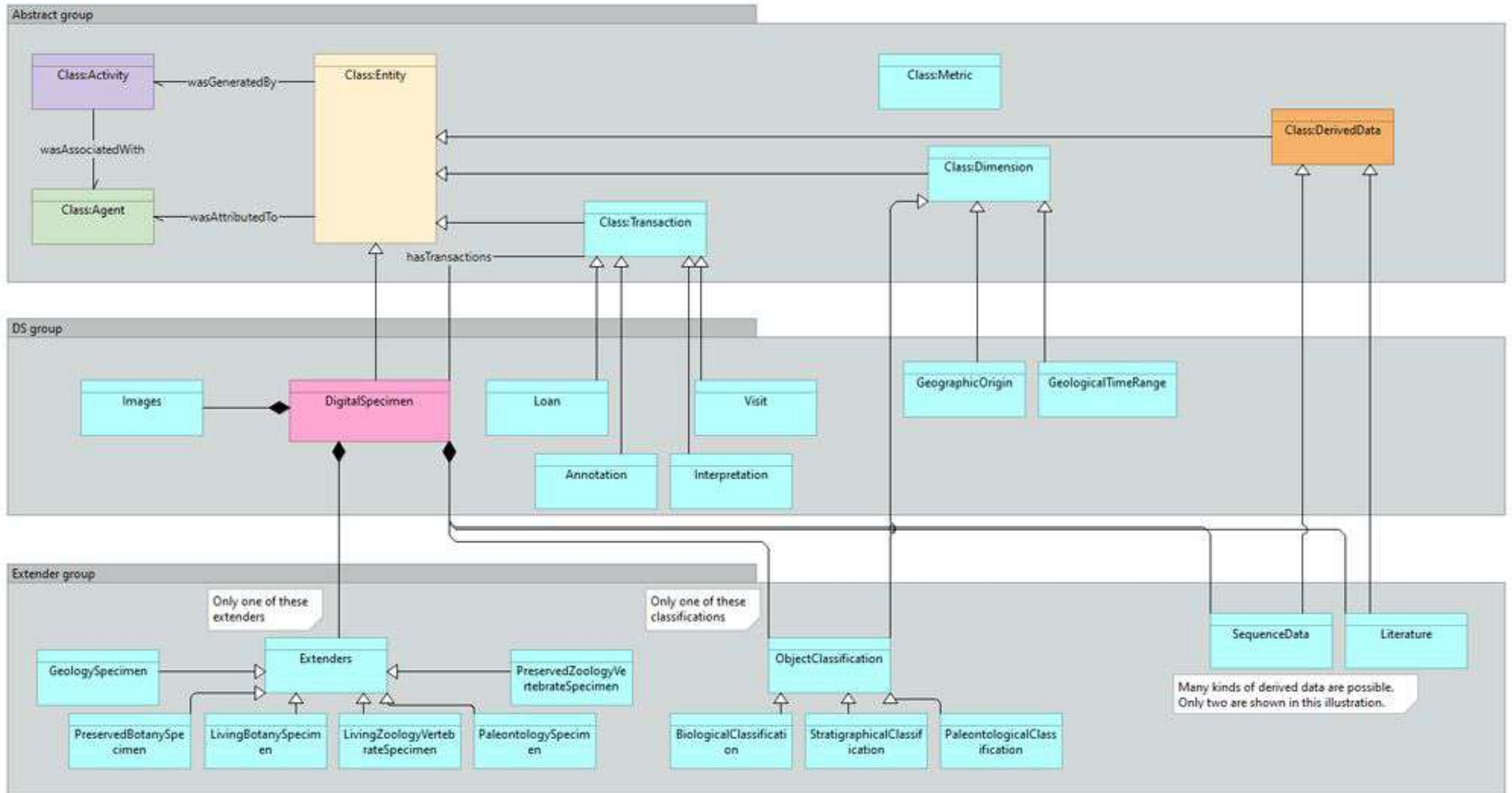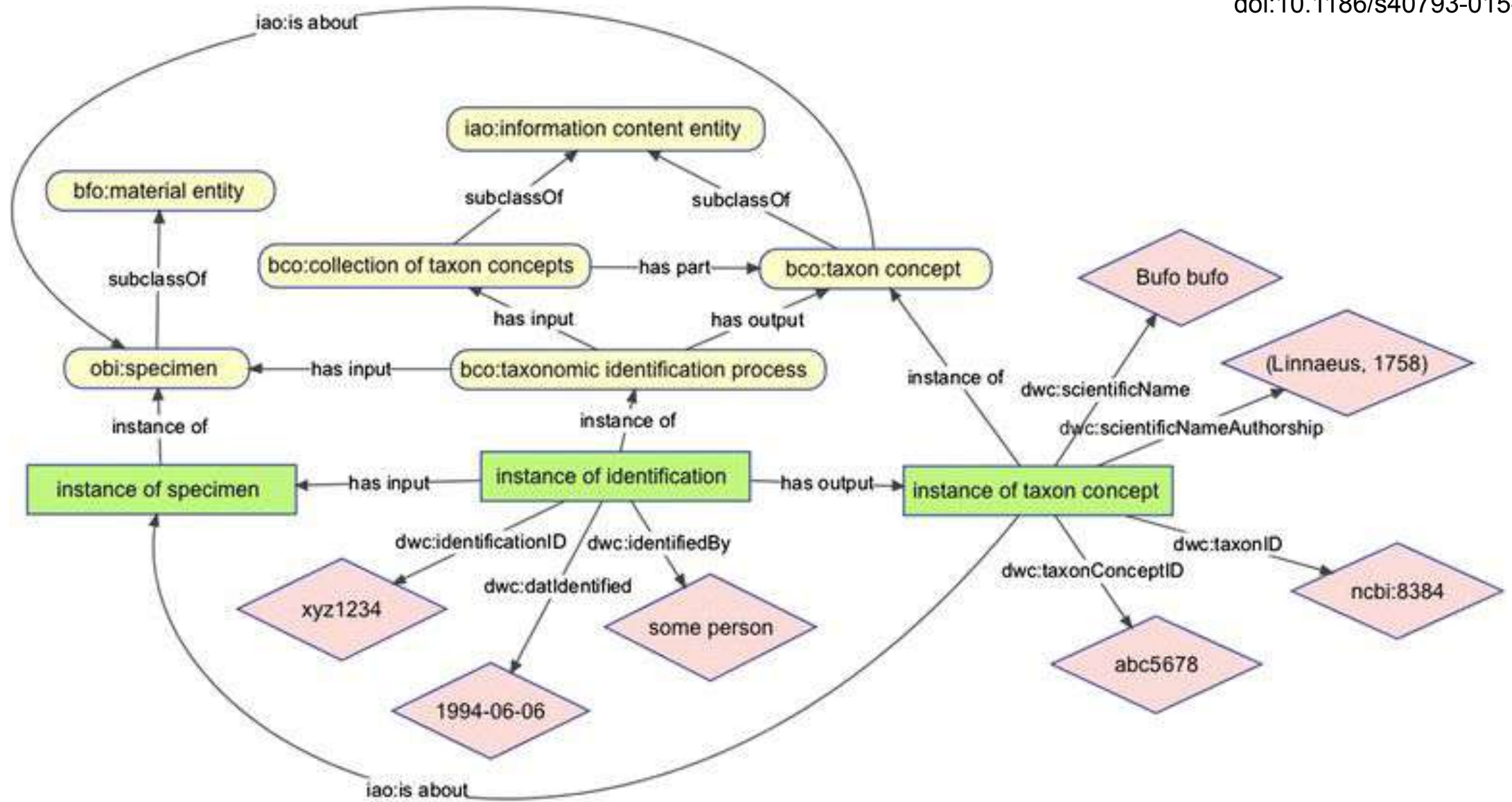
Ontological representation of the identification process in the Biological Collections Ontology

A few challenges

- CoL getting data from DiSSCo (like a reference to a type specimen or new names)
- Resolvable persistent identifiers for CoL names and taxon concepts
- Showing different opinions and names from different name providers (e.g. NCBI taxonomy)

- Resolvable persistent identifiers for GeoCase
- Requirement of BioCASe provider tool installations next to IPT
- Governance model for GeoCase

- Combined progress indicators for data shared with GBIF and GeoCase
- DiSSCo user stories that need implementation in GeoCase or CoL
- Digital Specimen data in GeoCase beyond ABCD/EFG

Integration of biodiversity data into EOSC through a flexible semantic mapping framework (SEMAF)

Claus Weiland

openDS II
DPP AHM1 (January 18-22 2021)

What is SEMAF about?

- EOSC funded study lead by CLARIN ERIC in cooperation with GEDE to support the implementation of EOSC by addressing cross-discipline semantic interoperability issues e.g. integration of data originating from other research communities.

- Interoperable exchange of such data is enabled through the use of ontologies that provide "meaning" to data and enable reasoning.

- Aim: Needs assessment for a service infrastructure to support registering, sharing and reuse of semantic interoperability solutions, in particular mappings and alignments/crosswalks of ontologies (Mapping = semantic relation of 2 entities from 2 ontologies (equivalence ($\equiv$), disjunction ($\perp$)), an alignment/crosswalk is a set of mappings).

- semantic mappings should be made machine-readable and machine-actionable, and should include state-of-the art existing data models, in particular FAIR Digital Objects.

DiSSCo PREPARE
Virtual meeting

Project state (modified slide from Broeder 2020)

- Design study:
  - Infrastructure & data model
    - Federation of registries
  - User interface
    - Easy specification of mapping relations
  - Machine access
    - Harvestable API
  - Operational & content management requirements
    - Bulk imports of existing alignments
  - Implementation requirements
    - PIDs for mappings and crosswalks to follow FDO

- Needs Assessment: 23 interviews with the following distribution: biology (3), biomedicine (2), environment (6), natural (3) humanities (4), social science (3), generic (1), incl. DiSSCo members from BGBM, MBG

SEMAF REGISTRY

crosswalk ← metadata ← resolves to ← PID

mapping ← metadata ← resolves to ← PID

DiSSCo
PREPARE
Virtual meeting
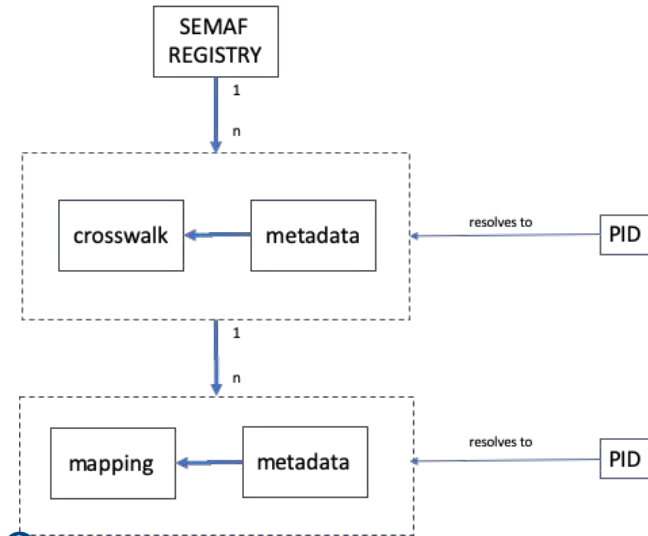
Why is this relevant for Biodiversity Sciences?

- Large datasets are provided by mass digitization efforts, which open up large amounts of trait and environmental data, linking these data to digital resources like sequence databases as well as the setup of biodiversity infrastructures like nsidr.org highlights the **need for bridging technologies to facilitate re-use of biodiversity data**.

- A strong focus of current ecological research lies on the l**ink between the environment and functional traits,** enabling the identification of consistent mappings between different trait and environmental ontologies (Karam et al 2020).

- To enable analysis and knowledge discovery of corresponding biodiversity knowledge graphs, machine learning methods based on **graph representation learning** were developed, e.g. demonstrating the usage of ontology meta-data and annotations (github.com/bio-ontology-research-group/machine-learning-with-ontologies).

DiSSCo
PREPARE
Virtual meeting

Semantic Mappings for Geobiodiversity Data Map DwC Location to OBO ontologies (BCO)

- Deck et al: Provide mappings between BCO with DwC, in particular import DwC basisOfRecord terms (classes of DwC) into BCO

- Example: Add an axiom to BCO that dwcobo:Location is equivalent to bfo:site, provide specification of Location as domain of geospatial DwC properties enabling usage from the OBO framework, here obi:specimen collection process

- "semantically light" OWL representation of DwC, template for OBO-adjacent efforts for other geobiodiversity semantic artifacts?

- github.com/BiodiversityOntologies/dwcobo
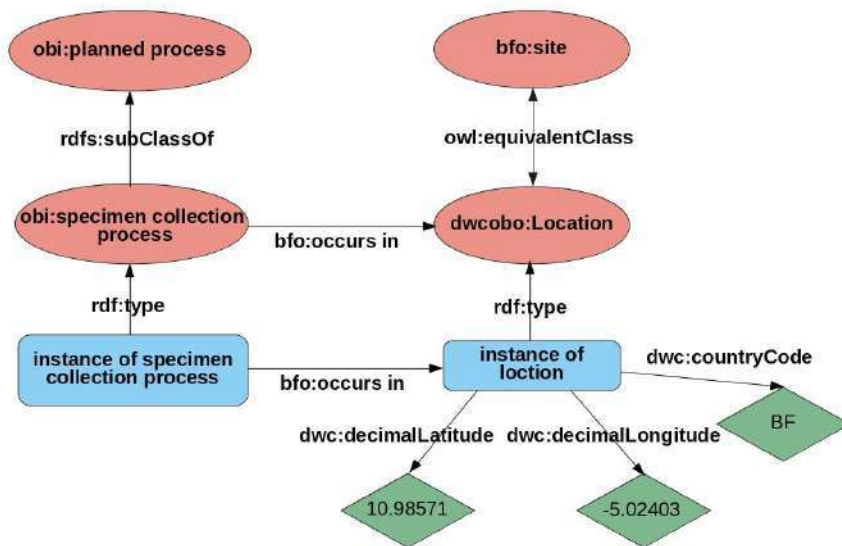
Modified Figure from Deck et al (2015) doi:10.1186/s40793-015-0014-0

Representing Taxonomic Identification including a Digital Specimen (again based on Deck et al 2015)
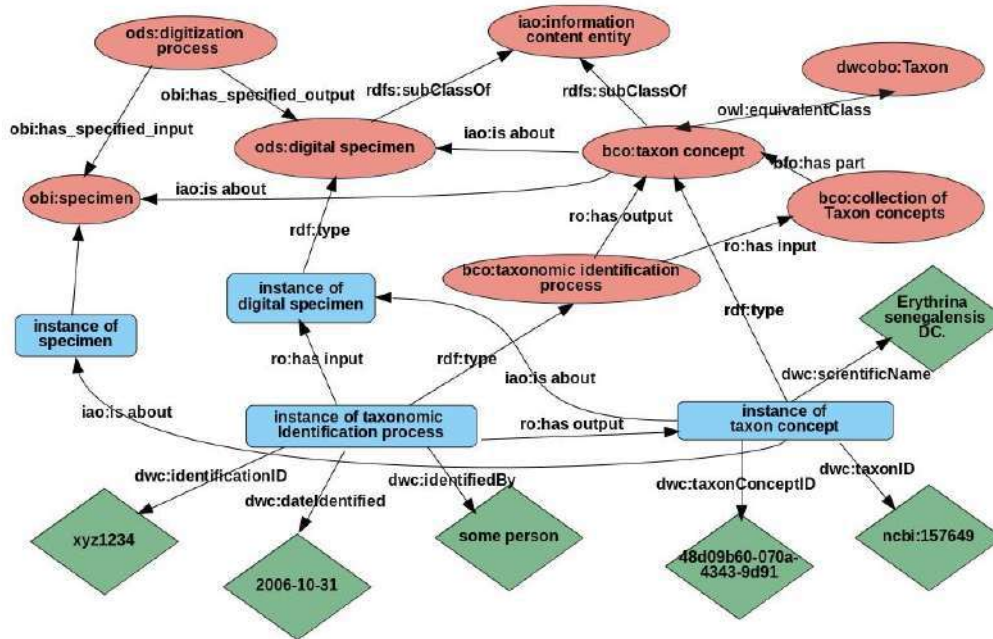
- Model of taxonomic identification process that gets information from a Digital Specimen and has as output a taxon concept that feeds back (iao:about) to Physical and Digital Specimen

- Enables reasoning to track the provenance of taxonomic identification processes and link specimens to a taxon concepts (here simplified, many in the original model)

- Scope: Only to demonstrate usefulness to build on and link to existing or currently developing semantic infrastructures like OBO, SEMAF, dwcobo approach etc etc

  ○

Modified Figure from Deck et al (2015) doi:10.1186/s40793-015-0014-0

Thanks!

- SEMAF task force: Daan Broeder, Paolo Budroni, Wim Hugo, Emiliano Degl'Innocenti, Keith Jeffery, Yann Le Franc, Claus Weiland, Peter Wittenburg, Carlo Maria Zwolf

- [SEMAF mission statement](#)

Intentional blank node ;-)

Neuro-symbolic gap between biodiversity KGs and machine learning

- **KGs rest on symbolic representations**, expressed particularly in ontologies e.g. GO (function of genes), ENVO (environment) & FLOPO (phenotypes of flowering plants).

- **Machine learning algorithms operate on numeric representations of data** (images, class labels etc) in n-dimensional vector space (**sub-symbolic feature vectors**)

```
Obo:FLOPO_0002102
    a owl:Class ;
    rdfs:label "fruit fleshy"^^xsd:string ;
    ...
    owl:intersectionOf (obo:PO_0009001
    [
    a owl:Restriction ;
    owl:onProperty obo:RO_0000053 ;
    owl:someValuesFrom obo:PATO_0002351
    ]
```
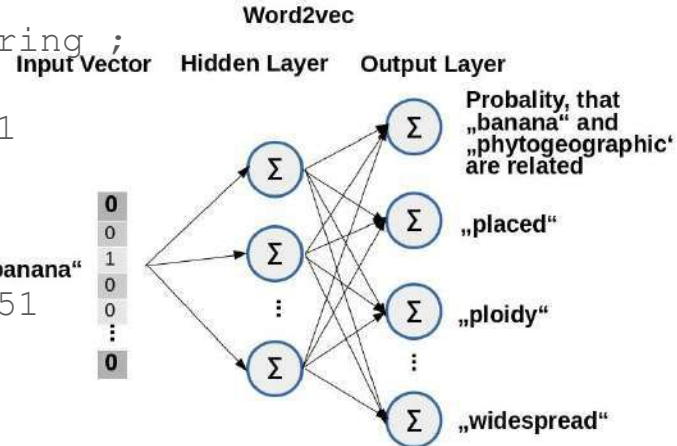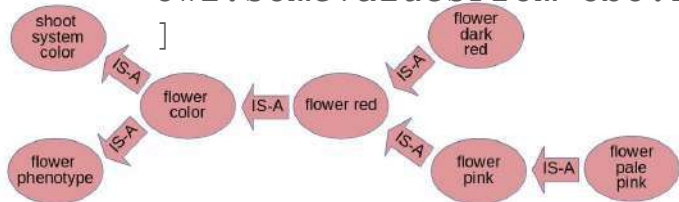
Extraction of trait data based on Deep Learning
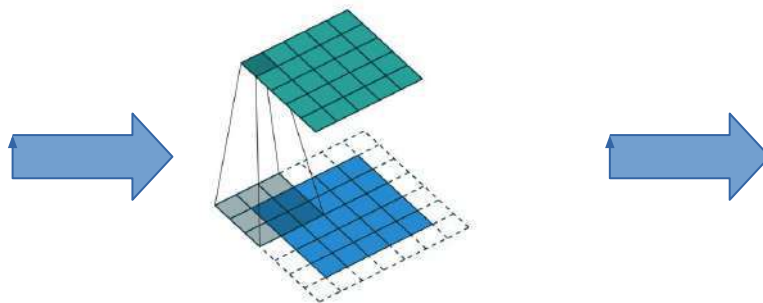
Modified ResNet18 (He et al. 2016)
(Animation by Paul-Louis Pröve 2017)
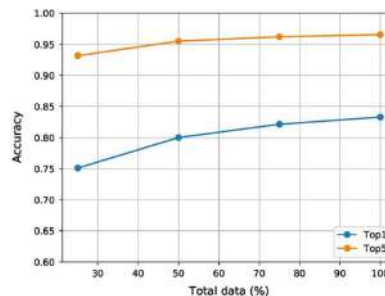
in depth PS0201

Younis et. al.

Employ Graph Convolutional Networks (GCNs)
- ## to improve classification & infer unseen graph properties

- Relational Graph Convolution Networks (R-GCNs) can be applied sucessfully to the multi-relational data of Kgs (Schlichtkrull et al 2017)

- R-GCN implementation as multigraph with entities as nodes and relations as labelled edges for node classification and link prediction (insert new triples).

$$h_i' = \sigma \left(\sum_{r=1}^{R} \sum_{j \in \mathcal{N}_i} \alpha_{ij}^r \mathbf{W}_r \mathbf{h}_j \right)$$

; where alpha is based on edge type
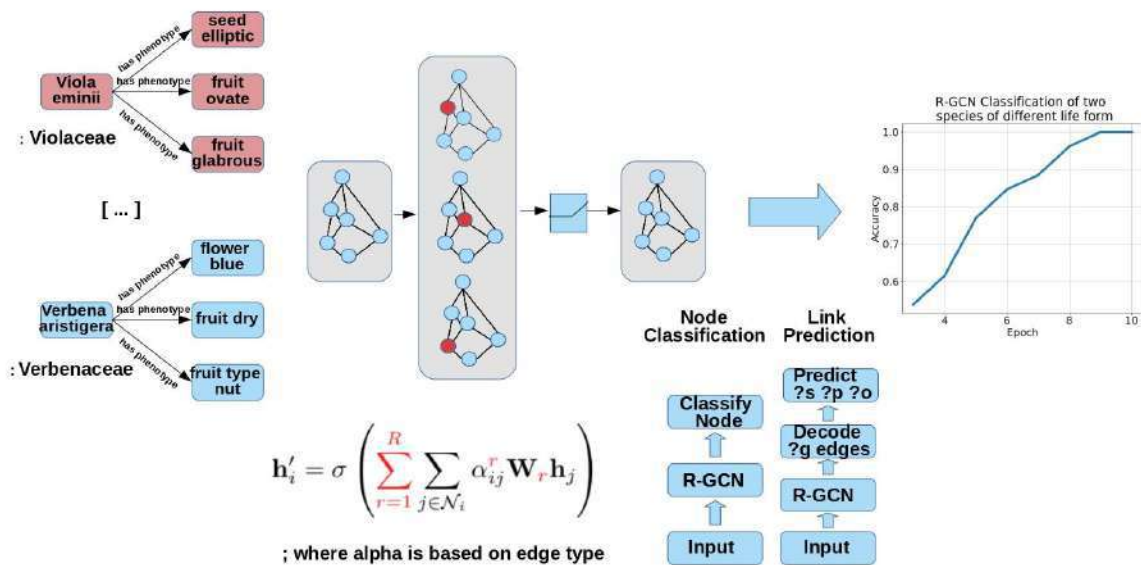
FAIR DIGITAL OBJECT AND RDA OUTPUT INCORPORATION IN DiSSCo's DESIGN

SHARIF ISLAM [0000-0001-8050-0299](ORCiD) @gnu111

Data Architect

DiSSCo / Naturalis Biodiversity Centre (Leiden, The Netherlands)

AHM, WP6 SESSIONS

https://dissco.eu

Hype Cycle for Emerging Technologies, 2020

Expectations (y-axis)

Time (x-axis)

Secure Access Service Edge (SASE)
Social Distancing Technologies
Embedded AI
Data Fabric
Explainable AI
Composable Enterprise
AI Augmented Development
Carbon-Based Transistors
Responsible AI
Multiexperience
Digital Twin of the Person
Citizen Twin
Packaged Business Capabilities
Generative AI
Composite AI
Adaptive ML
Bring Your Own Identity
Social Data
Private 5G
Generative Adversarial Networks
Differential Privacy
Biodegradable Sensors
Ontologies and Graphs
2-Way BMI (Brain Machine Interface)
Health Passport
Self-Supervised Learning
Low-Cost Single-Board Computers at the Edge
DNA Computing and Storage
AI-Assisted Design
Authenticated Provenance

| Innovation Trigger | Peak of Inflated Expectations | Trough of Disillusionment | Slope of Enlightenment | Plateau of Productivity |

Plateau will be reached:
○ less than 2 years   ● 2 to 5 years   ● 5 to 10 years   ▲ more than 10 years   ⊗ obsolete before plateau

As of July 2020

gartner.com/SmarterWithGartner

Gartner.

Software Development
Architecture and Design 2020 Q2 Graph

http://infoq.link/architecture-trends-2020

InfoQ

Innovators

Micro frontends

Data Mesh

AsyncAPI

Policy as Code

Blockchain

HTTP/3

RSocket & Reactive Streams

Early Adopters

Modular Monolith

Correctly built distributed systems

Actor Model

"Serverless"

Service Mesh

Functional programming

"Architect as technical leader"

workflow and decision automation platforms

Low code/no code

CHASM

Early Majority

GraphQL

Reactive programming

HTTP/2 and gRPC

Event-driven architecture

CQRS

Event sourcing

Eventual consistency

Late Majority

Microservices

Domain-driven Design

Behavior-driven Design

Test-driven design

REST

Gergely Orosz
@GergelyOrosz

For the record, at Uber, we're moving many of our microservices to what @copyconstruct calls macroservices (wells-sized services).

Exactly b/c testing and maintaining thou microservices is not only hard - it can ca trouble long-term than it solves the shor

> **Cindy Sridharan** @copyconstruct · Apr 6, 2020
>
> - Microservices are hard.
> - Building reliable and testable microservices is a lot hard think
> - Effectively *testing* microservices requires a ton of too
> - A Netflix/Uber style microservices isn't required by mar
> - Macroservices?
>
> Show this thread

2:03 PM · Apr 6, 2020 · Twitter Web App

changelog.com/posts

Monoliths are the future

by Kelsey Hightower · 2020-01-30 · #practices +2

What you're about to read is an excerpt from Go Time #114. You should read the original transcript or listen to the entire conversation for more context. To set the stage, this is the 2nd of Kelsey Hightower's unpopular opinions (our new segment which people seem to be enjoying.)

Monoliths are the future because the problem people are trying to solve with microservices doesn't really line up with reality. Just to be honest - and I've done this before, gone from microservices to monoliths and back again. Both directions.

FAIR DIGITAL OBJECT
(historical context)

- Digital Object: "A digital object is a data structure whose principal components are digital material, or data, plus a unique identifier for this material, called a handle" – Kahn and Wilensky, 1995 (http://www.cnri.reston.va.us/k-w.html).

- FAIR (Findable, Accessible, Interoperable, Reusable) - Wilkinson et al. 2016.

- "Digital Objects have the built-in capability to implement FAIR data" – George Strawn, 2019.

- FAIR Digital Object "..is a stable actionable unit that bundles sufficient information to allow the reliable interpretation and processing of the data contained in it" - De Smedt, K., Koureas, D. and Wittenburg, P., 2020.

- https://www.go-fair.org/today/fair-digital-framework/

Cornerstone: Digital Object (DO)

An important tenet of our approach is to move from a system-centric world to an information-centric world.

Corporation for National Research Initiatives

Reference: Giridhar Manepalli (CNRI) 2019
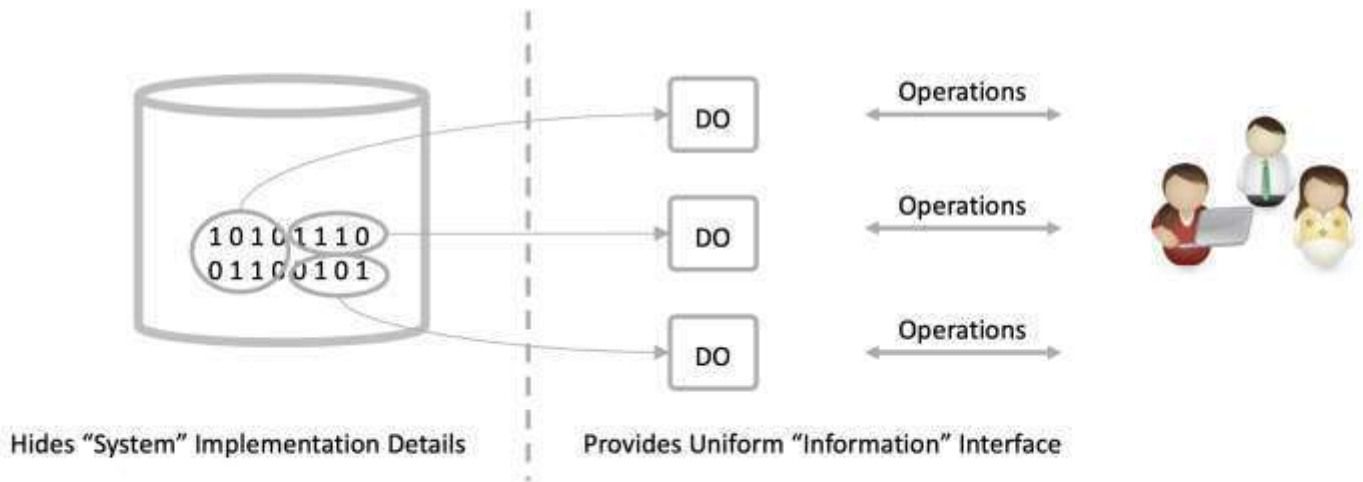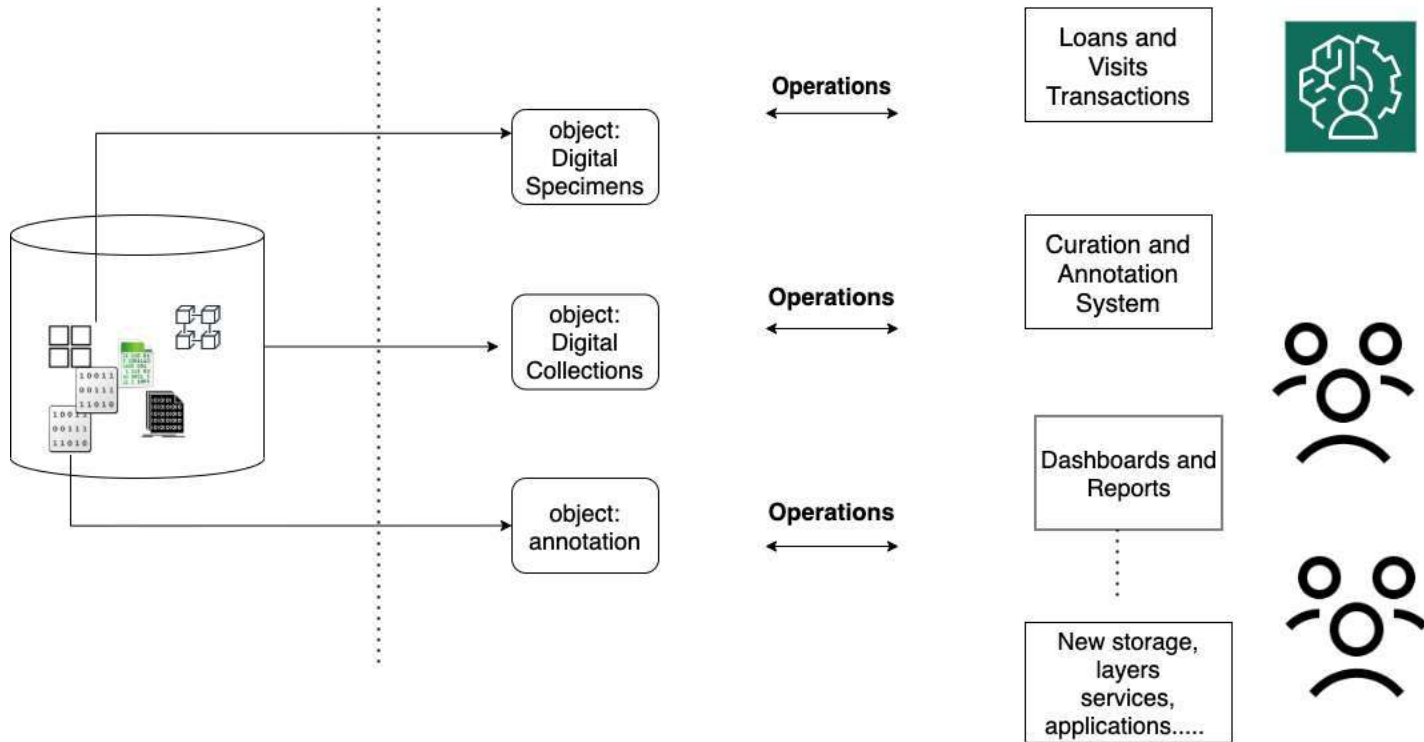
A Digital Object oriented approach for Natural Science Collections data:
Vision for an endless number of levels of abstraction

Why FAIR Digital Objects for natural science collections data

- To use heterogenous data sources

- Need digital representations of physical specimens – unambiguous, persistent

- Different levels of abstraction

- Different actors and agents

- Collect and anchor core information about the specimen

- Information needs to be persistently and unambiguously linked to necessary context for interpretation and validation

- Standard and domain specific operations for specific applications

We are building "systems" and "services"

- System wide perspective on complexity

  - Local complexity: Depends on the implementation of a service. Example: taxonomic name resolution, PID resolution.

  - Global complexity: Interaction and dependencies between the services. Example: How CMS interacts with a PID system.

- We cannot build systems out of independent components.

- Various services will always have to interact with each other to form a (**distributed**) system (**of scientific collections**).

- *Reference: Myers, G. J. (1978). Composite/structured design. New York: Van Nostrand Reinhold. https://catalog.hathitrust.org/Record/010130734*

TECHNICAL AND COMMUNITY ALIGNMENT

Recommendations   Guiding Principles
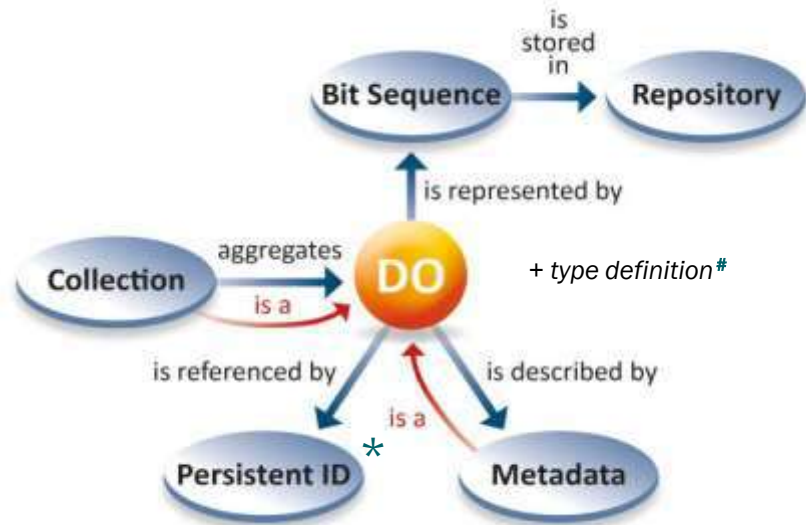
*Resolvable via well-known resolver,
#Accessible via a type registry

DO = Digital Object

DiSSCo Building Blocks

DiSSCo Design Decisions and FAIR data lifecycle

	Community Resource	Workflow	DiSSCo Element
Adoption of Digital Object Architecture	RDA Data Foundation and Terminology WG Data Fabric and Terminology IG	All phases of the data lifecycle	Digital Specimen Architecture
Persistent Identifiers and Kernel Information	RDA PID Kernel WG		- Meta-information about a digital object - DiSSCo (data) type registry

- FAIR data principles
- Global discussions around FAIR Digital Objects
- Outputs from the Research Data Alliance (RDA) interest and working groups
- ENVRI FAIR, GOFAIR Initiative

DiSSCo Design Decisions and FAIR data lifecycle

	Community Resource	Workflow	DiSSCo Element
Aggregation of digital objects	RDA Research Data Collection WG	All phases of the data lifecycle	Digital Specimen Architecture
Metadata attribution and use of PROV entities	RDA/TDWG Metadata attribution WG	Digitization, curation and maintenance of digital object	Digital Specimen and collection objects
FAIR data maturity model	RDA FAIR data maturity model WG, ENVRI FAIR, GOFAIR Initiative	All phases of the data lifecycle	Develop guidelines and specifications to assess FAIR implementation plan

STRUCTURE OF A DIGITAL SPECIMEN DIGITAL OBJECT (DSDO)
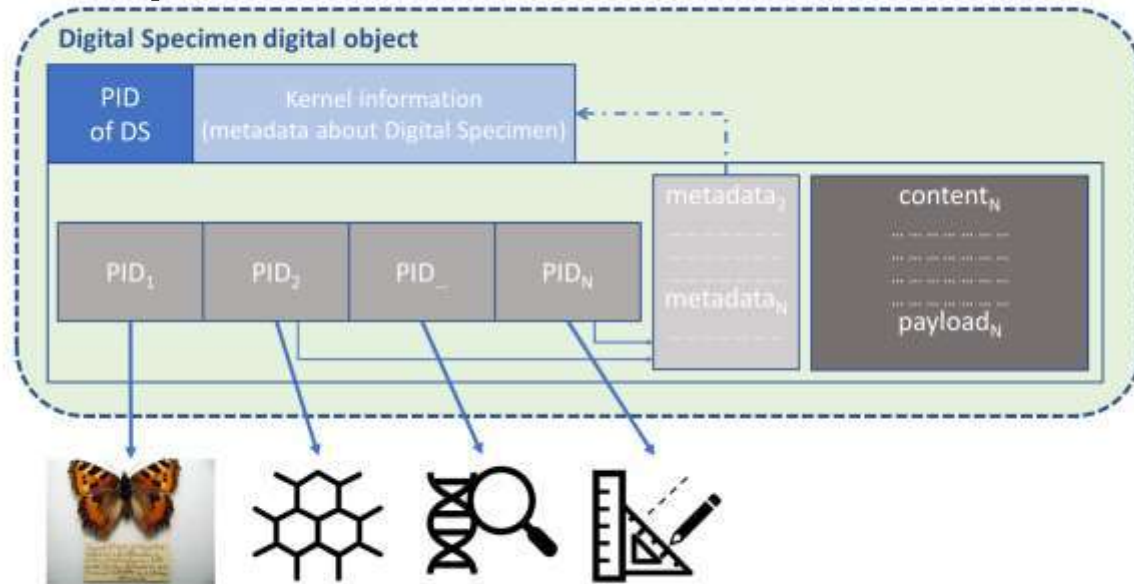
- PIDs are pointers that resolve to location (URL) of the item e.g., DO itself, physical specimen, hi-res images, label information, tissue sample, DNA sequence, etc. Image credit: Alex Hardisty.

A simple Digital Specimen
Image Credit: Alex Hardisty, Feb 2020

ICEDIG.EU

doi: 10.11646/zootaxa.1426.1.3   Publications

BMNH:2006.12.6.4041

hasCircumscription   has

alsoKnownAs   hasHolotype

hasSequence

Physical specimen
BMNH:2006.12.6.40-41

MNHN JNC1848-D1

isParatypeOf
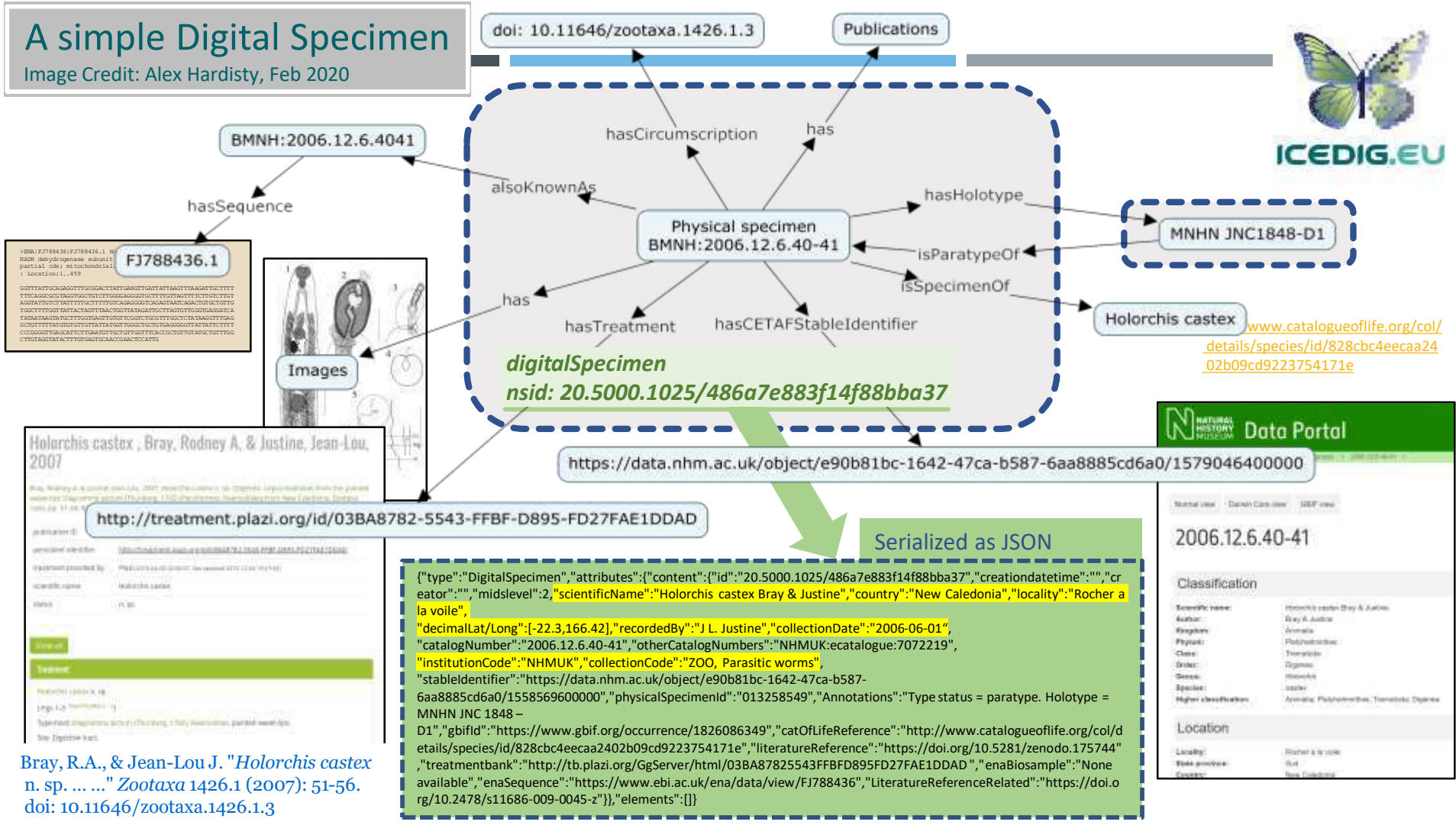
FJ788436.1

>ENA|FJ788436|FJ788436.1 H...
NADH dehydrogenase subunit
partial cds; mitochondrial
; Location:1..459

isSpecimenOf

has   hasTreatment   hasCETAFStableIdentifier

Holorchis castex

www.catalogueoflife.org/col/
details/species/id/828cbc4eecaa24
02b09cd9223754171e

Images

digitalSpecimen
nsid: 20.5000.1025/486a7e883f14f88bba37

https://data.nhm.ac.uk/object/e90b81bc-1642-47ca-b587-6aa8885cd6a0/1579046400000
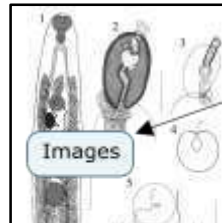
Holorchis castex , Bray, Rodney A. & Justine, Jean-Lou, 2007

NATURAL HISTORY MUSEUM Data Portal

2006.12.6.40-41

Classification

Location

http://treatment.plazi.org/id/03BA8782-5543-FFBF-D895-FD27FAE1DDAD

Serialized as JSON

{"type":"DigitalSpecimen","attributes":{"content":{"id":"20.5000.1025/486a7e883f14f88bba37","creationdatetime":"","creator":"","midlevel":2,"scientificName":"Holorchis castex Bray & Justine","country":"New Caledonia","locality":"Rocher a la voile",
"decimalLat/Long":[-22.3,166.42],"recordedBy":"J L. Justine","collectionDate":"2006-06-01",
"catalogNumber":"2006.12.6.40-41","otherCatalogNumbers":"NHMUK:ecatalogue:7072219",
"institutionCode":"NHMUK","collectionCode":"ZOO, Parasitic worms",
"stableIdentifier":"https://data.nhm.ac.uk/object/e90b81bc-1642-47ca-b587-
6aa8885cd6a0/1558569600000","physicalSpecimenId":"013258549","Annotations":"Type status = paratype. Holotype =
MNHN JNC 1848 –
D1","gbifId":"https://www.gbif.org/occurrence/1826086349","catOfLifeReference":"http://www.catalogueoflife.org/col/d
etails/species/id/828cbc4eecaa2402b09cd9223754171e","literatureReference":"https://doi.org/10.5281/zenodo.175744"
,"treatmentbank":"http://tb.plazi.org/GgServer/html/03BA87825543FFBFD895FD27FAE1DDAD ","enaBiosample":"None
available","enaSequence":"https://www.ebi.ac.uk/ena/data/view/FJ788436","LiteratureReferenceRelated":"https://doi.o
rg/10.2478/s11686-009-0045-z"}},"elements":[]}

Bray, R.A., & Jean-Lou J. "*Holorchis castex*
n. sp. ... ..." *Zootaxa* 1426.1 (2007): 51-56.
doi: 10.11646/zootaxa.1426.1.3

DiSSCo Building Blocks

Incorporating RDA Outputs

- **Recent paper:** Islam, S., Hardisty, A., Addink, W., Weiland, C. and Glöckler, F., 2020. Incorporating RDA Outputs in the Design of a European Research Infrastructure for Natural Science Collections. Data Science Journal, 19(1), p.50. DOI: http://doi.org/10.5334/dsj-2020-050

Incorporating RDA Outputs

Simple example of PID Kernel Information for a Digital Specimen. Example PID: 123prefix/uuid-27a9edf63.

Attribute	Value Type	Example Value
Location	url	http://example-dissco-repo/uuid-27a9edf63
Created	date and time	2019-04-24T11:07:11.771Z
Type	type definition	typedef123/DigitalSpecimen
PhysicalSpecimenId	string	BMNH:1905.5.30.352

Source: http://doi.org/10.5334/dsj-2020-050

ACKNOWLEDGEMENTS

- Questions?

Biohackathon 2020

33: Connecting molecular sequences to their voucher specimens

Quentin Groom, Mathias Dillen, Pieter Huybrechts, Maarten Trekels -- Meise Botanic Garden, BE

Rukaya Johaadien -- Natural History Museum, Oslo, NO

Niki Kyriakopoulou -- Naturalis, NL

Francisco Quevedo -- Cardiff University, UK

Wai Yee Wong -- University of Vienna, AT
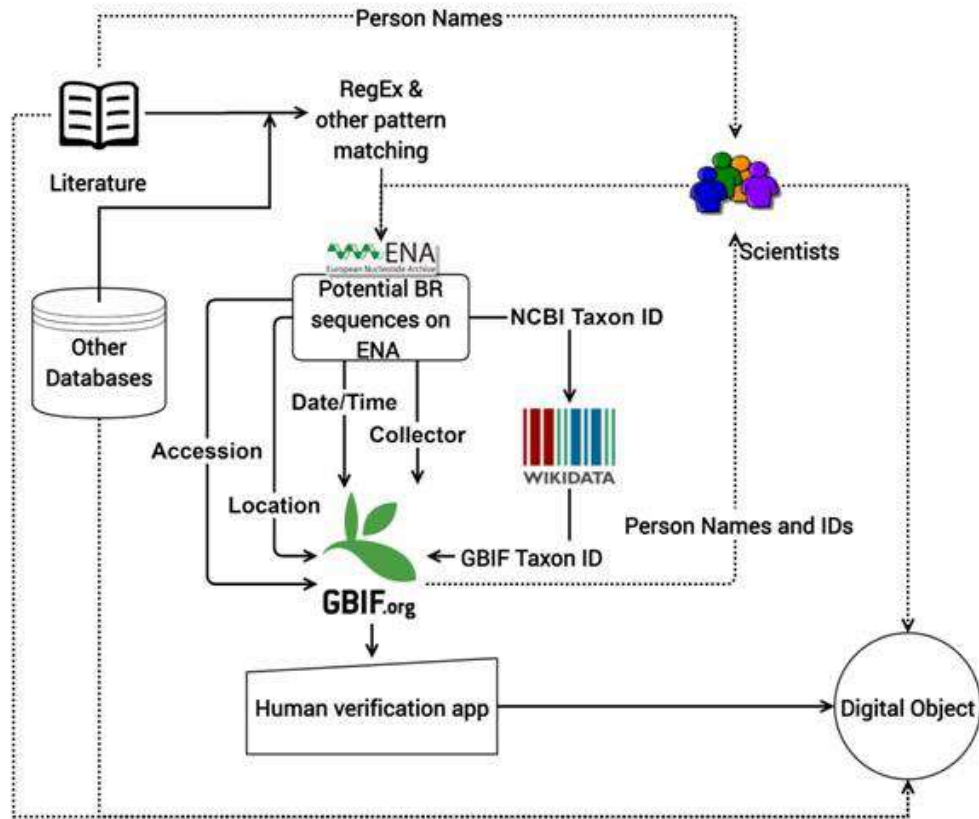
ENA: European Nucleotide Archive; Part of **INSDC** (International Nucleotide Sequence Database Collaboration)

Finding MeiseBG sequences

- Query the ENA sequences API
- Linking MeiseBG specimens
  - Specimen_voucher property:
    - Specimen identifiers?
    - DWC triplets?
      - APM:BR:BR0000025959222V

- In practice: IH code (BR), (old) institution names

```
*br)*

*br:*

*br<*

*br-*

*meise*

*gard.*belg*
```

Match to specimens (GBIF)

- specimen_voucher: often consists of
  - Institutional identifier (i.e. BR)
  - Person name
  - Numeric specimen/accession identifier

- Matching process based on:
  - Collector
  - Taxon id
  - Collector number

examples:

Lemaire & Verstraete 142 (BR)

BR:De Block:694

BR-Stoffelen et al. 13

BR<BEL>:De Block et al. 2294

Match to specimens (GBIF): collector

- MeiseBG publishes gbif:recordedByID
- Also maintains list with (sur)names and their associated PIDs

- Match known surnames into specimen_voucher string
- List positive matches by their recordedByID

Match to specimens (GBIF): taxon ID

- ENA provides a NCBI taxon ID
- Wikidata as a broker to connect NCBI taxon ID to GBIF backbone ID
- Find all links with series of SPARQL queries:

```
SELECT ?taxon ?taxonLabel ?ncbi_taxonID ?gbifid WHERE {
        VALUES ?ncbi_taxonID {"642057" "659064" "703364" "884208" "884211" "884213" "91480" "947892" "980240" "980244" }
        ?taxon wdt:P685 ?ncbi_taxonID.
        OPTIONAL {?taxon wdt:P846 ?gbifid .}
        SERVICE wikibase:label { bd:serviceParam wikibase:language "[AUTO_LANGUAGE],en". }
    }
```

NCBI: National Center for Biotechnology Information

Match to specimens (GBIF): numbers

- Botanical specimens often have collector numbers to identify the collection event
  - Complicated by lack of uniformity
- Multiple numbers may be present in specimen_voucher

- Extract space-separated numerics from specimen_voucher
- See if any match to numerics on GBIF

Initial results

- 1.336 out of 5.920 found sequences enriched (23%)
  - Some ambiguously
  - Set up a Django (Python) app for human-in-the-loop disambiguation

DO integration

- DOIP update operation of existing specimen record
- nsidr.org demonstrator: e.g. http://nsidr.org/#objects/20.5000.1025/ad7de78d504378fcec03

DO integration

- DOIP update operation of existing specimen record
- nsidr.org demonstrator
  - Add generic supplementary info:
    - Identifier
    - Identifiertype
    - relation

  - Specific ENA/INSDC property

```
"$id": "#/properties/suppinfo/link",
"type": "array",
"title": "External Link",
"items": [
  {
    "type": "string",
    "title": "Domain"
  },
  {
    "type": "string",
    "title": "Relation"
  },
  {
    "type": "string",
    "title": "IdentifierType"
  },
  {
    "type": "string",
    "format": "uri",
    "title": "Identifier"
  }
]
```

```
"enaSequence": {
  "maxLength": 128,
  "format": "uri",
  "type": "string",
  "$ref": "#/definitions/previewInResultsFalse",
  "title": "Sequence data (European Nucleotide Archive)"
},
```

Further refinement & discussion

- Living accessions
  - And their connections to preserved specimens
- Truncated identifiers
- Mining literature
- Mining lab logs?

Recommendations

- ## For INSDC partner databases

  - Databases should incorporate PIDs into their data model for people, institutions, publications and specimens.

- ## For collections

  - Scientists depositing sequence data should be given training on the data model and standards used.

  - Collections should make more effort to reconnect their backlog of voucher specimens to their sequences.

  - Specimens should obtain a PID at the earliest point possible upon collection and certainly before tissue collection for sequencing.

Documentation and code

R and Python

https://github.com/elixir-europe/BioHackathon-projects-2020/tree/master/projects/33

BioRxiv report in preparation.

DPP All Hands Meeting 1

T7.1 Governance structure – Core elements and landscape analysis - Brainstorming session

Eva M. Alonso.

DiSSCo Prepare Project Manager

Task 7.1 Leader – Refinement of the governance model, strategy & operational planning

Funded by the Horizon 2020
Framework Programme
of the European Union

Agenda

Time	Topic	Lead
15.00-15.10	Welcome & Introduction	EA
15.10-15.20	Rationale & initial work. • Purpose of the session • Initial outcomes • Objective	EA
15.20-15.50	1. Best practices: requirements Outcomes: List of requirements & RIs	All
15.50-16.10	2. Core element Outcome: List of core elements	All
16.10-16.25	Timeline & Next steps	All
16.25-16.30	AoB	All

Purpose & expected outcomes of the session

Purpose of session

This session aims to
1. identify core elements in the governance structure proposed.
2. set up requirements to analyse operational Ris and their governance models in other RIs.

Both elements contribute to the preparation of MS7.4 *update/upgrade of the governance model* and the process to set up DiSSCo LE.

Expected outcomes

1. A list of core elements of the governance chart,
2. List of requirements and RIs (best practices),
3. Agreement on the timeline.

Initial outcomes

✓ DiSSCo EU MoU
       = Signing process

✓ Existing interim governance chart and core elements in the actual description of the interim gov. model (EU MoU)

✓ Coordination with DiSSCo Aspiration WG
                  = To refine DiSSCo mission/vision

✓ Close cooperation with T7.2 & WP8
                  = Alignment with LE & Consultation process

Objective
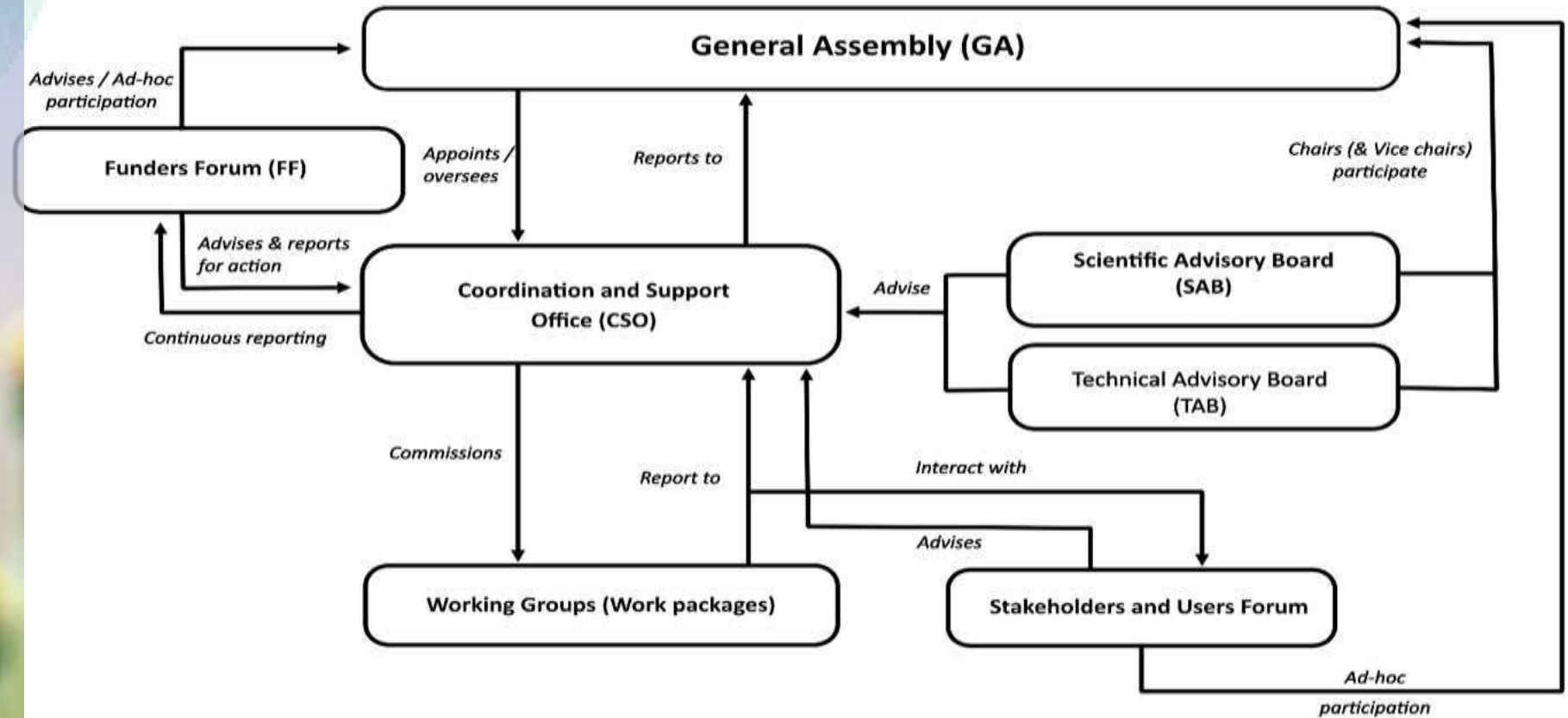
(D7.1) **A detailed plan for governance structure and function, and participation framework**
- Due: Jan 31st 2022 - Extended Feb 28th 2022

Starting point: DiSSCo specificities

1	What are the core elements we need to preserve?
2	Which provisions will guarantee long-term sustainability (both, financial and service provision)?
3	What will be the contribution model? How in-kind contributions will be articulated?
	How institutions participate meaningfully in the decision-making process? and in the review/upgrade of services and policies? At what level do we see their participation?
4	What type of institutional relations will be set up with DiSSCo? How cultural differences will be articulated?
5	How the technical community of DiSSCo also drive the RI developments? How should we reconcile the technical needs and other community needs? How does it affect the contribution model?
6	Facilitate communication and strategic decisions between policy-makers and the IT-makers. How to do it?
7	What is the role of the international actors?
8	How DiSSCo will govern services? How international relations are articulated?
9	Functional and/or divisional? Flat or what degree of hierarchy?
10	...

Starting point: Landscape analysis

	CRITERIA FOR A LANDSCAPE ANALYSIS
1	Country-based RI
2	Community rooted - important institutional role
3	Technological challenge
4	Purpose/Mission
5	Distributed data-driven RI
6	… other criteria coming from the main questions/DiSSCo specificities ...

Timeline

MS/D	ACTIONS	LEAD	Jan W3	Jan W4	Feb W1	Feb W2	Feb W3	Feb W4	Mar W1	Mar W2	Mar W3	Mar W4	Apr W1	Apr W2	Apr W3	Apr W4	May W1	May W2	May W3	May W4	Jun W1	Jun W2
	1. Refinement of the gov.model & feasibility analysis of its implementation	**Naturalis**																				
Ms7.2	EU MoU update (Ms7.2). Consultation process with NNs and part of supporting documentation GA2																					
	2. Define the future governance (incl. ToR) of the DiSSCo legal entity																					
Ms7.4	Proposal for an update/upgraded gov. model (due date anticipated: 30/11/2021)																					
	Workshops to discuss and prepare 1st draft gov model	Naturalis/MNHN/RBINS	AHM1			M		M		M												
	Workshops for contributions	All										M	M									
	Consultation process with NNs	CETAF												NN								
	Preparation of updated/upgraded gov. model	Naturalis/MNHN/RBINS																				
	FF Assessment (FF extraordinary meeting)																		FF			
	GA3 Approval																					GA
	Workshops to discuss and prepare 1st draft ToRs	Naturalis/MNHN/RBINS																				
	Workshops for contributions	All																				
	Consultation process with NNs	CETAF																				
	Finalisation of the ToRs	Naturalis																				
	3. Define DiSSCo RI Strategy and Strategic Planning																					
Ms7.3	Series of workshops to prepare the Draft Strategy & operation planning (WG will deliver before June the mission and vission)																					
	Naturalis starts 1st draft of Stragegy and operational planning for discussion + Workshop with WP4	Naturalis/All									WP4											
	Workshop for contributions	All												M								
	WG delivers mission and vission (tbd)	WG												?	?	?	?					
	Consultation process with NNs	CETAF													NN							
	Finalisation of DiSSCo Strategy and operational planning framework (only)	Naturalis																				
	FF assessment																		FF			
	GA3 Approval																					GA

"Bringing the irreplaceable data stored in natural science collections to life

and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

The preparatory phase project of DiSSCo Research Infrastructure
Distributed System of Scientific Collections

https://www.dissco.eu/prepare/

All Hands Meeting 1

Presentation of the study on the Legal Entity model(s) proposed and discussion with Legal Advisor as guest

January 20th 2021

DiSSCo Prepare WP7 T7.2

Session "Towards a Legal Entity Model for DiSSCo"

Royal Belgian Institute of Natural Sciences

Agenda of the session

Time	Description	Speaker
15h00 – 15h10	Welcome & Presentation of the task and session	Carole Paleco, RBINS
15H10– 15h20	Overview of the work and process leading to the Milestone of T7.2	Serge Scory, RBINS
15h20 – 15h30	Analysis of the LE models and recommendation for DiSSCo	Ohad Graber-Soudry, Xofficio
15h30 – 15h40	CETAF key role in DiSSCo RI	Ana Casino, CETAF
15h40 – 15h50	EOSC RI as an AISBL	Patricia Mergen, MBG
15h50 – 16h00	EMBRC ERIC model and Relations with the French government	Vanessa Demanoff, François Dusoulier, MNHN
16h00 – 16h10	Experience with ERIC Forum	Eva Alonso, Naturalis
16h10 – 16h25	Q&A on key aspects and feedback from participants	All
16h25 – 16h30	Wrap up & Conclusions	Serge Scory, Carole Paleco

Link to the programme on Teamworks

https://dissco.teamwork.com/#/files/9236593

Note takers:
Carole Paleco, Vanessa Demanoff, François Dusoulier

TASK 7.2: Towards the creation of a legal entity

PARTNERS: RBINS, Naturalis, MNHN, BG Meise, CETAF + WP7 NHM, LUOMUS

With this task, we will prepare the necessary steps to establish DiSSCo as a legal entity in order for the RI to deliver its services in the most efficient way. The following actions will be performed in sequence:

- Detailed analysis of the legal entity models, including a SWOT analysis, and their suitability for achieving DiSSCo objectives in agreement with its governance model. The result of the analysis will be presented and **discussed with the various national contact entities (cf. WP 8), in order to reach a consensus on the best choice.** Draft statutes and by-laws will be prepared, according to the legal entity model that appears to be the most appropriate.

- Planning for the creation of the legal entity.

TASK 7.2: Towards the creation of a legal entity

PARTNERS: RBINS, Naturalis, MNHN, BG Meise, CETAF + WP7 NHM, LUOMUS

Milestone

MS 7.1 - Analysis of the legal entity models and their suitability for achieving DiSSCo objectives (report)
Submitted: January 15th 2021   = Synthesis of a two-phase detailed analysis of 6 models and recommendation to opt for an ERIC

Deliverable

D 7.2 – Draft statutes and by-laws; implementation plan
Due: December 31st 2021   = Operationalization of the chosen legal entity model

Next steps

We will have to work on the hypothesis the consultation of the stakeholders and the decision by the DiSSCo General Assembly confirms the recommendation
Decision expected after the Summer holidays

A quick question on your experience with Legal Entity models !
Go to slido.com and enter the following code event :

Overview of the work and process leading to the Milestone of T7.2

Milestone MS 7.1 - Analysis of the legal entity models and their suitability for achieving DiSSCo objectives (report) Submitted: January 15th 2021

- Screening of the resource documentation
  - Icedig Blueprint D8.1 & D9.4
  - ESFRI Strategy report on Research Infrastructures Roadmap 2018
  - Examples from other RIs - Lifewatch, eLTER, EOSC, EMBRC, ...
- DiSSCo legal basis, the MoU to provide the objectives and governance suited for the LE of DISSCo. The MoU has been updated under T7.1.

- RI landscape analysis complex and need for an expertise to understand the models, their characteristics, specificities and assets appropriate to DiSSCo

Overview of the work and process leading to the Milestone of T7.2

Icedig Conceptual design blueprint for the DiSSCo digitization infrastructure
DELIVERABLE D8.1 - and 6. Governance - "DiSSCo Preparatory phase governance"

&

Deliverable 9.4 Positioning DiSSCo among other research infrastructures
notably : positioning with CETAF, EOSC, GBIF…

Data available on https://icedig.eu/content/research-infrastructure-introduction

Screening of the landscape through key analysis reports from the ESFRI

The Landscape Analysis provides the current context of the most relevant Research Infrastructures that are available to European scientists and to technology developers typically through peer review of competitive proposals. The unique contribution played by the ESFRI RIs in all scientific domains is analysed along with the interconnections and cross-cutting aspects of the whole European RIs ecosystem.

http://roadmap2018.esfri.eu/strategy-report/

Strategy Report on Research Infrastructures
ROADMAP 2018

PART2
LANDSCAPE ANALYSIS

ENVIRONMENT

ACTRIS
Aerosols, Clouds and Trace gases Research Infrastructure

ENE — Aerosol, clouds, and trace gases observations for energy supply and consumption
Key atmospheric parameters and environmental assessments related to different energy production forms

ENE — Aerosol, clouds, and trace gases observations for energy supply
Diagnostics development for wind and solar radiation energy applications

H&F — Aerosol, clouds, and trace gases observations for health and food
Measurements of air quality parameters and impact on health
Measurements of air quality parameters and extreme weather events and impact on agriculture

PSE — New technologies and approaches for aerosol, clouds, and trace gases observations
Development of laser optics and detectors
New techniques for atmospheric corrections and aerosol physics

SCI — Social and cultural aspects about aerosol, clouds, and trace gases observations
Key atmospheric parameters, climate change and impact on human wellbeing

DIGIT — Open data resources and management for aerosol, clouds, and trace gases observations
Data access and curation, standardisation of data ad metadata, data and services interoperability

DIGIT — ICT tools for atmospheric research
Virtual Research Environments (VRE) for data processing, trend analysis, and satellite cal/val
HPC for atmospheric global and regional models, data assimilation and processing of large volume of data

DANUBIUS-RI
International Centre for Advanced Studies on River-Sea Systems

H&F — Advanced studies on river-sea systems for health and food
Surface waters analysis in global biogeochemical cycles, food and energy production, food security, aquaculture, environmental medicine

SCI — Sustainability scenario development for human societies in river-sea systems
Integration of data from interdisciplinary Earth and Life sciences with social, economic ad behaviour information on communities living in river-sea systems, aiming to support sustainable management plans at entire basin scale

DIGIT — Open data resources and management for interdisciplinary research on river-sea systems
Development of DANUBIUS Commons – common set of standards, rules, methodologies supporting interdisciplinary research in riverine, transitional and coastal marine environments

DiSSCo
Distributed System of Scientific Collections

H&F — System of scientific collections for health and food
Natural science collections for bio- and geo-diversity information, climate change, food security, health and bioeconomy

DIGIT — Open data resources and management for scientific collections
Multi-modal access to collections for a linked open data approach

EISCAT_3D
Next generation European Incoherent Scatter radar system

PSE — New technologies and approaches for monitoring of the atmosphere and ionosphere
Development of new methods of radar coding, signal processing and data analysis

PSE — New technologies and approaches for monitoring of the atmosphere and ionosphere
Sun-Earth interactions, radio astronomy, space security, plasma physics, magnetic field studies

DIGIT — Open data resources and management for monitoring of the atmosphere and ionosphere
Collection and archive of basic data; data processing to describe the ionosphere and neutral atmosphere; selection of well-designed radar pulse schemes

Screening of the landscape through key analysis reports from the ESFRI

The Landscape Analysis interconnexions between RIs fields

& Looking for the RIs LE models chosen among the Landmarks
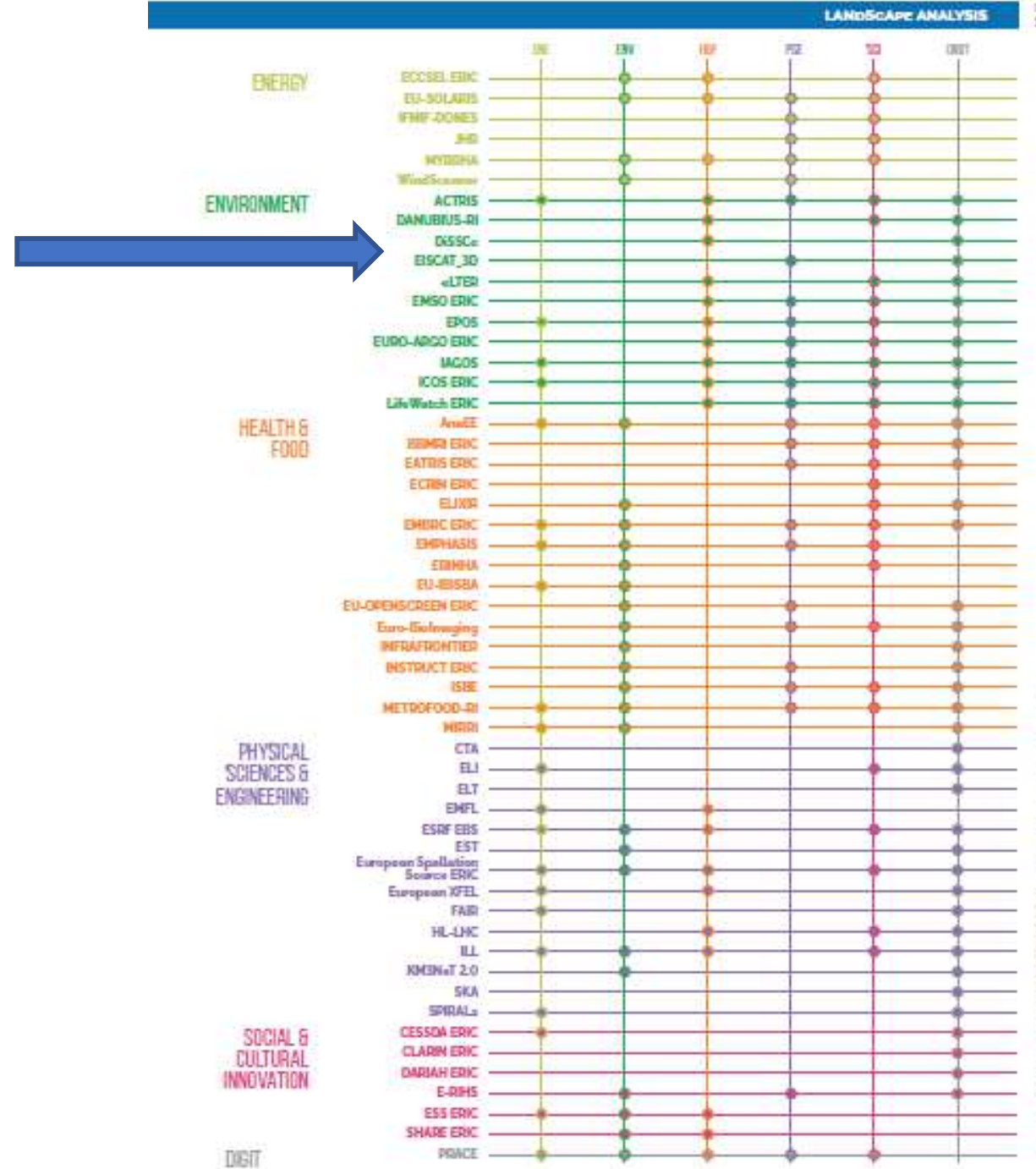
Definition of the initial questions/criterion with the partners

- Institutions & organizations and governments must have their say, scope of the voting right/membership;

- Legal capacity recognized in all participating countries: does the LE model allow the same level of engagement whatever the country the partners/members belong to ?;

- How simple and fast is the establishment process?;

- What is the procurement regime (including VAT and excise regime)?;

- Does the LE model allow non-for-profit commercial activities?

- Are there initial capital requirements for incorporation?

- What is the liability regime for the legal entity itself, the governing body and for the members?

- How flexible is the governance structure?

- Ability to receive EU and national grants, to contract with public and private third parties; hire personnel, buy equipment

- Ability to get bank loans

LE models

Analysis

Short list

Two-step analysis performed by our legal advisor

LE models

Dr Ohad Graber-Soudry

X-officio
LEGAL | GOVERNANCE | PROCUREMENT

X-officio provides legal, governance and procurement support to inter-governmental and national research infrastructures, ERICs, research facilities and laboratories used by the science community to conduct research and foster innovation.

X-officio supports departments of administration, COOs, DG/CEOs and national ministries in all matters pertaining to legal, governance and procurement in the process of establishing, constructing and operating a research infrastructure. We also support business and contractual partners in legal and procurement matters.

Analysis

Short list

Conclusion ← **Analysis** ←

Legal Entity for DiSSCo

DiSSCo Prepare WP7 T7.2

All Hands Meeting | January 2021 | Ohad Graber-Soudry | www.xofficio.eu

Background

X-officio commissioned by RBINS to prepare a report on the most suitable long-term legal entity for DiSSCo

Phase one:

- Six optional legal entities

- 15 assessment criteria

- Shortlisting

Phase two:

- Three optional legal entities

- Practical assessment

- Recommendation

Legal entities considered

- International/Intergovernmental Organisation (IO)

- European Research Infrastructure Consortium (ERIC)

- European Economic Interest Grouping (EEIG)

- European Groupings of Territorial Cooperation (EGTC)

- Belgian AISBL

- Dutch Stichting

Main recommendations

The AISBL and ERIC are the most suitable long-term legal entities for DiSSCo.

ERIC is preferable over the AISBL in a number of important aspects (provided there is sufficient support by the relevant ministries).

A number of mitigation steps could be implemented in order to address specific concerns related to the participation of CETAF in the General Assembly and the decision-making process of the ERIC.

Thank You

Ohad Graber-Soudry
ohad.graber-soudry@xofficio.eu
+46 769 435368
www.xofficio.eu
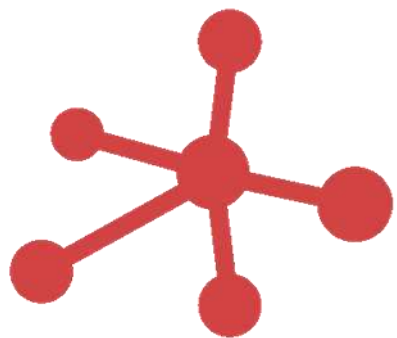
X-officio

LEGAL | GOVERNANCE | PROCUREMENT
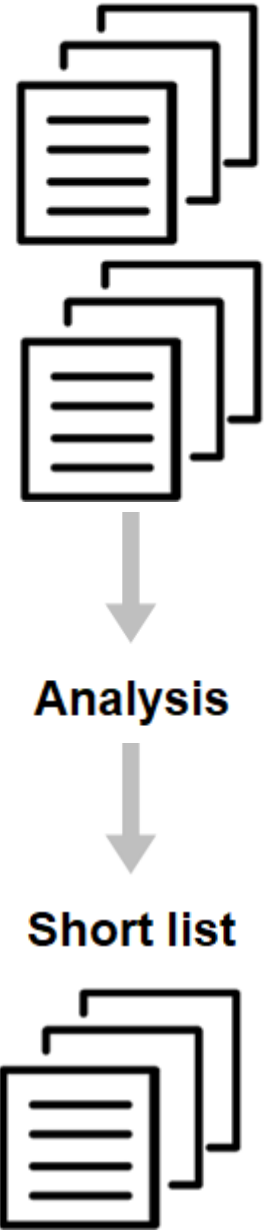
Final assessment

	AISBL	ERIC
Level of integration between the DiSSCo legal entity and the national nodes	+	++
Location of the statutory seat	+	++
Language to be used for the official documents	+	++
Membership	Institutions, individuals	Countries, IOs
Participation by CETAF	++	+
"Branding" and network possibilities	+	++
Procurement and VAT considerations	+/–	++
Tailor-made governance structure	+	++
Securing funding	+/–	++

Recommendation

LE models

On the basis of the analysis by the legal advisor, and the discussions with him and the partners associated with the task, we recommend to proceed to the next steps with the working hypothesis that establishing DiSSCo as an ERIC is the best option, taking into account that adequate mechanisms need to be implemented to allow a prominent role for CETAF.

Analysis

Short list

Recommendation ← Assessment ← Conclusion ← Analysis ←
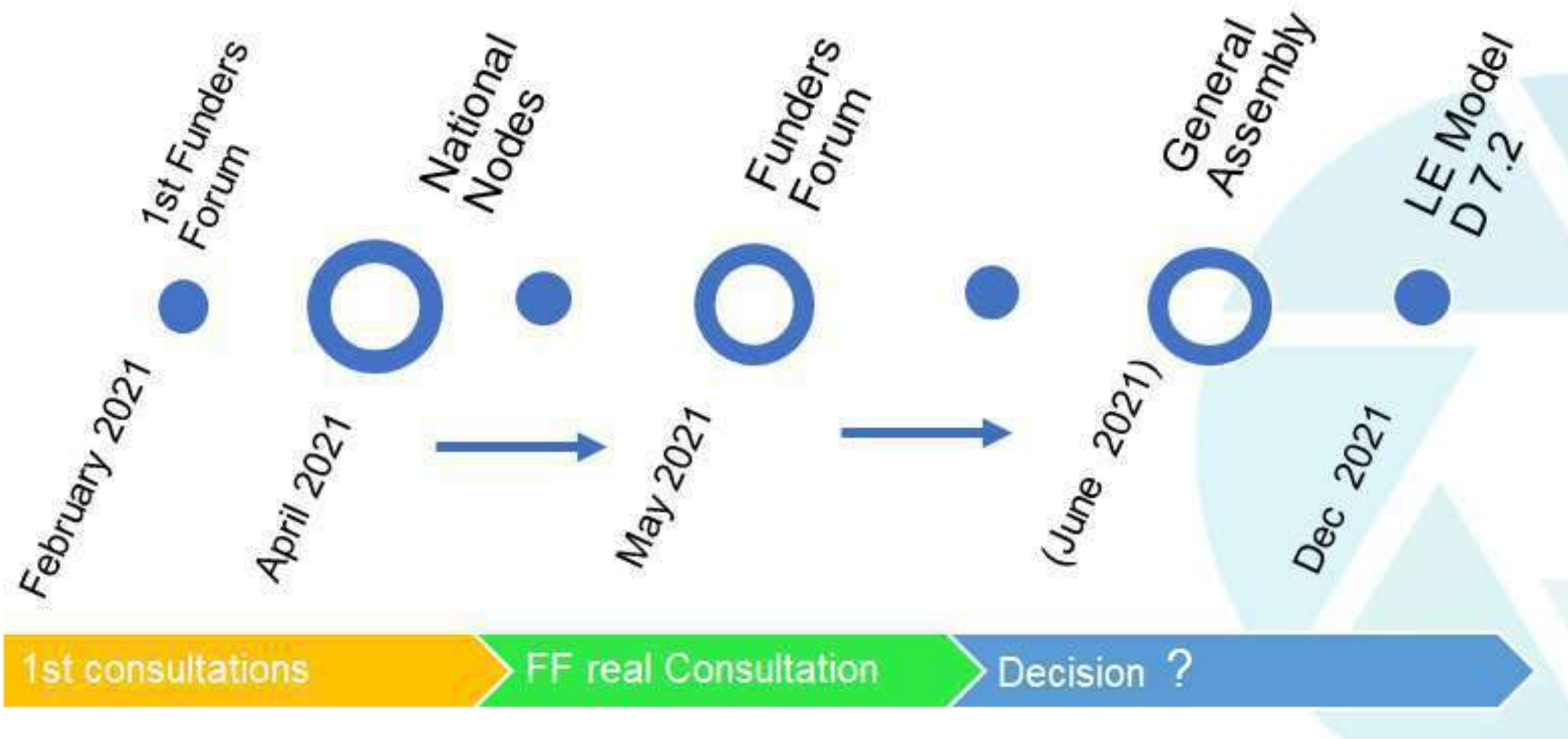
ALL HANDS MEETING - AHM1

WP7- Governance, Policy & Legal frameworks

18 January 2021

Ana Casino - CETAF

DiSSCo Prepare

T7.2 – Legal Entity model

A bit of history

Starting back in 2013

- **Regular submission of proposals**…
  but most **failed!**
  Green Science (2013), Roadmap (2013), BioUnify (2014), SPECIFY (2015), EUColl (2015), DEDDI (2016), CELLS (2016), SYNYO (2016) …..

- Others, such as EUColCOMP (2013-15), SYNTHESYS3 (2014-17) and EUBON (2013-17), still **succeeded!!**

- **Frustration** feeling:
    - Lack of calls dedicated to the NSCs community
    - Collections out of focus
    - Need for thoughts out-of-the-box

Time passes…
…a renewed approach was shaped…

Initial conversations started in 2015

- **Small team** (DK-NHM London, MH-Naturalis, AC-CETAF) shared high-level vision to start working on
  - Learning from past mistakes
  - Leveraging SYNTHESYS funds towards building a ESFRI proposal (Nov 30, 2015)
  - KoM Naturalis Biodiversity Center, Leiden, NL (March 11, 2016)

- Support of **7 (largest) institutions and CETAF** to move forward, providing financial support and knowledge from previous initiatives

- An entire **community behind** thanks to CETAF involvement

Proposal for a Research Infrastructure

….the big opportunity arrived…

Submission to ESFRI 2018 Roadmap update (30 Aug'17)
Hearings (31 Jan'2018) (EvH, IO, DK)

After:
- Dozens of meetings
- Hundreds of contributors
- Innumerable discussions
- Countless versions of documents
- Multitude of events

Distributed System of Scientific Collections

….and we succeeded!!!

Proposal approval received July 16th, 2018 !
(Letter from Giorgio Rossi ESFRI Chair)

The process continues

with CETAF involvement in the supporting process

1) Plethora of linked projects under the DiSSCo overall programme
In which CETAF participates and leads Streams, WPs and Tasks
- ICEDIG
- MOBILISE
- SYNTHESYS+
- DiSSCo Prep
- ENVRI-FAIR
- CoL+

Design study

Blueprint

Networking

Access development

Maturity roadmap

DiSSCo Masterplan

Participation of CETAF

….and being essential partner to DiSSCo development

2) DiSSCo Coordination Office
- Deputy Coordinator (ExD)
- Communications and Engagement (CSO-CE) stream leader (ExD)
- Member of the Coordination Steering Board (7+1 entities) (Chair)

3) DiSSCo interim Governing body
- Full member of the DiSSCo iGA (since 2019) (ViceChair & Secretary)

4) Thematic synchronization WGs
w/ CETAF WGs

5) SAP and advisory bodies and technical teams
w/ coordinators

Role of CETAF – since the beginning

The DiSSCo approach

- **Transfer of authority** from facilities to central for all key operations
- Clear **decision making** mandate
- Focused **scientific scope**
- **Binding institutional commitments** (already in place)
- New **independent legal framework**
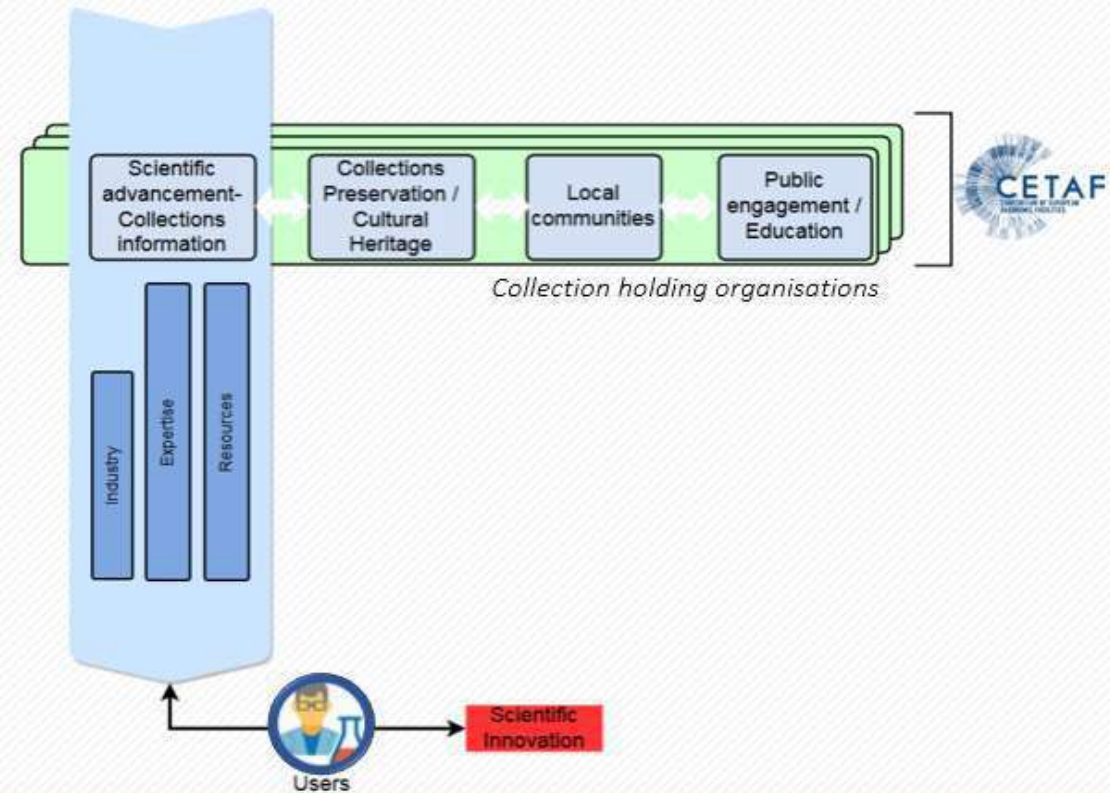
Q1

- **CETAF**: A 20-year network acting as a forum for discussion and advocacy
- **DiSSCo**: An integrated research infrastructure providing data, services and links to other infrastructures

The community effort at the core

*"The partnership of DiSSCo worked together for a number of decades and **organised its cooperation in CETAF.** They produced, sometimes as output of collaborative projects, a substantive number of reports in which the design of DiSSCo components is elaborated"*

*"Since a number of decades, a growing number of European natural science collections investigated and materialised several modes of cooperation ....(...).Together, this **cumulative experience provides a robust corpus of technical, governance and socio-cultural feasibility reports and studies** that drives the development and operation of the new Research Infrastructure. A driving force is the strong, stable and cohesive Consortium of CETAF."*

*"The Consortium of European Taxonomic Facilities (CETAF), the cooperative body of distributed collection facilities, was **instrumental in developing a programme of joint research and development** projects."*

Scientific case for the DiSSCo Design study.

AH1 / WP7 – Task T7.2

CETAF
Policy areas and Working groups

- Collections ⟶ Registry of Collections
- Digitization ⟶ MIDS
- Earth Sciences ⟶ GeoCASE
- E-publishing ⟶ E-publishing guidelines / EJT
- Information Science and Technology (ISTC) ⟶ CETAF Stable Identifiers
- Legislation and Regulations ⟶ CoC and BP on ABS
- European Initiatives advisory group ⟶ RRI principles
- Biodiversity Monitoring ⟶ COST action
- Training and e-learning ⟶ DEST
- Communication ⟶ Communication material
- Scientific Research ⟶ Joint Research Agenda

AH1 / WP7 – Task T7.2

Critical added value to DiSSCo

DATA providers
- Needs and requirement channels
- Quality checks
- Accuracy and update (also for CSs data)
- Standardization
- Annotations and enhancement
- Consolidation and aggregation
- Peer-review
- Publication and mobilisation

SERVICES users

- Test-bed implementation
- Analysis and Review
- Model integration and linkage
- Reporting and knowledge merging
- Innovative research
- Multidisciplinar science
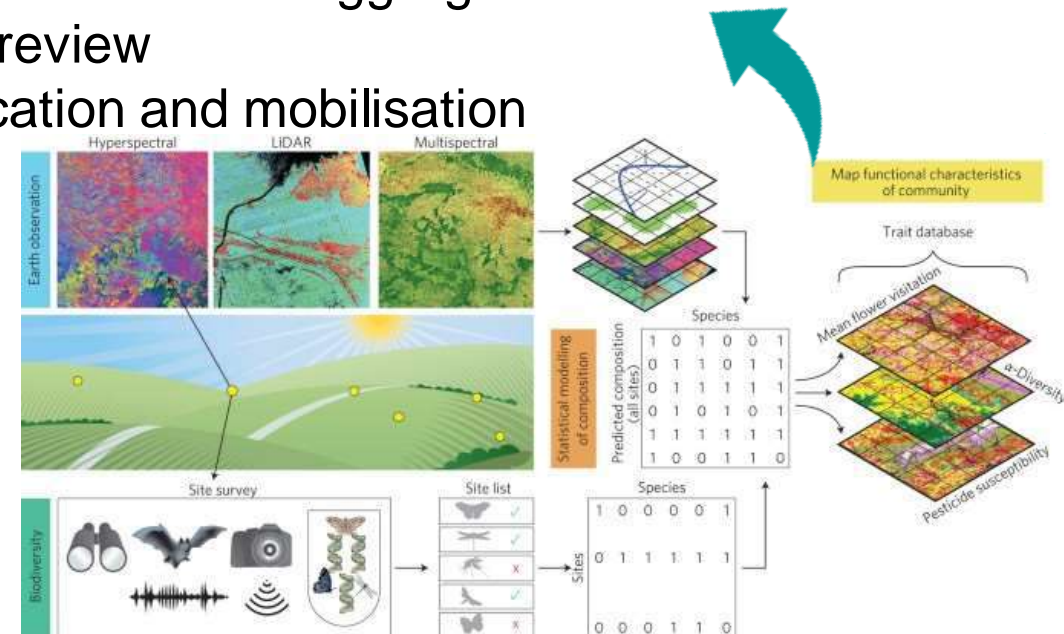
CETAF in DiSSCo

In the foreseen legal & governance framework….
Towards the upcoming RI building phases

Timing:
- After Preparatory phase is completed
- Before DiSSCo becomes fully operational

Funding:
- No direct EU support will follow for the RI construction
- Country level commitments will be required

Governance:
- To secure institutional representation
- To give voice to the community

ROLE to play / partnering model

From now onwards….

Considerations:

- Role as Provider of services (on top of data):
    - Expertise and training
    - Registry of collections
    - Scientific advise
    - Others….
- Profile identification
    - Participation of the institutions
    - Participation of the CETAF as clustering organization
- Integration in the DiSSCo Legal & Business model
    - Voting rights in governing bodies
    - Decision-making process
    - Cash / in-kind

To move forward, the participation of CETAF ensures DiSSCo RI

Integral: essential and fundamental, necessary to make a whole complete

Inclusive: comprehensive, towards equality, to avoid marginality and include minorities

Innovative: original and creative in thinking AH1 / WP7 – Task T7.2

Thanks for your attention!

Any question **?**

Task 7.2 DiSSCo Legal Entity
EOSC RI example as an AISBL

20° January 2021

Patricia Mergen - Meise Botanic Garden, Belgium

DiSSCo Prepare WP7

Governance, Policy & Legal frameworks

Funded by the Horizon 2020
Framework Programme
of the European Union

European Open Science Cloud (EOSC)

- Environment for hosting and processing research data

- Federated environment across borders and scientific disciplines

- Store, share, process and re-use research digital objects

- FAIR principles (Findable, Accessible, Interoperable, Reusable)

Recommended read:

A FAIR Lady : Solutions for a sustainable EOSC

https://www.eoscsecretariat.eu/

What legal model for EOSC ?

- EU Research Infrastructure supported by European Countries

- Scope European Continent beyond Member and AC States of the EU

- Not part of the ESFRI Roadmap process

- Should serve all possible users (EU bodies themselves, international bodies, other RIs, national bodies, private sectors …) of all sizes, even is not legal entities

Assigned to the EOSC WG Sustainability, Task Force Legal Entities

More info on the EOSC Working Groups : https://eoscsecretariat.eu/eosc-working-groups

What legal model for EOSC ?

- Study combining funding scheme and appropriate legal form

- Funding : Co-Programmed European Partnership under Horizon Europe.

Between the Commission and private and/or public partners. They are based on memoranda of understanding and/or contractual arrangements.

- Country support and decision making

Apply to ESFRI Roadmap future calls and become ERIC

- Issue ！ Deadlines of Horizon Europe and applying for Partnership incompatible

EOSC partnership proposal : https://ec.europa.eu/info/horizon-europe/european-partnerships-horizon-europe/candidates-across-themes_en

EOSC AISBL in practice

- Time constrains : Agreement to go for an AISBL

- Keep it open to apply for ESFRI and become an ERIC later

- AISBL set up end 2020

- Ready for the Partnership proposal and SRIA drafting

- Statutes and how to become a member available

- Currently working on the bylaws

- Over 100 members still growing off all sizes and disciplines including from countries that have not committed to support EOSC (yet)

More info here https://www.eoscsecretariat.eu/application-joining-eosc-association

EOSC AISBL in practice

- Member States not directly members. But in an external Steering committee ▯ *concern that advisory role, while many expect decision making roles*

- Mandated institution to vote on "strategic points" for their country ▯ *How to define what are strategic points ?*

- Regular Members to vote on the other points ▯ *What are the "other points"*

- Observer members with no votes

- How to serve all users and stakeholders that are not member of the AISBL as expected at country levels ?

EOSC AISBL in practice

- Faith and continuation of the EOSC Working Groups ⬚ Concern if the experts nominated by the Member States can remain in the Working Commissions of the AISBL, especially if they are not member of the AISBL ?

- Drafting of the Bylaws quite complicated

- Concern for the transition between the end "preparation phase" and time for the AISBL to be up and running that could take quite a long time.

EOSC material for DiSSCo

- EOSC produced many documents and studies useful for DiSSCo

- RIs important component of EOSC landscape analysis

- Agreement to collaborate and re-use outputs of EOSC

- Published and to DiSSCo potentially relevant documents can be found here (regularly updated):

https://drive.google.com/drive/u/0/folders/1atFsxjQ5WT7TG2fGewN6vvsDT-uVQBf

Any Questions ?

Meise Botanic Garden

DiSSCo PREPARE

HOW TO PLANT SOLID ROOTS FOR EOSC

3rd face-to-face meeting of the Sustainability WG

EOSCSustainability
Executive Board Working Group

February 2020

DiSSCo Prepare, AHM

Contact : Patricia.Mergen@plantentuinmeise.be

"Bringing the irreplaceable data stored in natural science collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

The preparatory phase project of DiSSCo Research Infrastructure - Distributed System of Scientific Collections

https://www.dissco.eu/prepare/

Task 7.2 DiSSCo Legal Entity
Feedback on EMBRC ERIC process

20th January 2021

Vanessa Demanoff, François Dusoulier
Muséum national d'histoire naturelle

DiSSCo Prepare WP7

Towards a Legal Entity Model for DiSSCo

The EMBRC ERIC Story

- Roscoff 8/9 october 2020

  - Bernard Kloareg – Director of Roscoff Marine Station / Director of EMBRC France

  - Anne-Emmanuelle Kervella (legal advisor) – Roscoff Marin Station / EMBRC France

  - Anne Nivart (MNHN / Ministry for Research)

  - Vanessa Demanoff (MNHN)

The EMBRC ERIC Story

➤ 1st step: Before choosing to become an ERIC

- Convince the communities and the researchers
- Define the ambition of the infrastructure:

  - service providing infra for ourselves (community) or is their a wider scope to consider

  - What road to follow: network, a service provider or Knowledge and Innovation communities (KIC)?

  Option 1: Research oriented (based on coordination by a European node, but relies strongly on the work and functioning of the national nodes)

  Option 2: Service oriented (delegate a number of decisions to a European node, which are no longer the competencies of national nodes and there is a common policy for the provision of services).

  ☐ Option 2 which is the most ambition and integrative option.

The EMBRC ERIC Story

➢ 2nd step: go for an ERIC straight away or start with a lighter structure first? ☐ ERIC

➢ End of 1st prep phase: EMBRC had a governance model to negotiate and discuss the statutes, and sign an MOU.

  • Negotiation and implementation of the governance was driven by the host country which therefore had to be chosen beforehand.

  • At this stage the support and implication of ministries / funders is essential.

➢ Phase 2: clarify budget, membership fees, integration of regions, etc.,

    ☐ several months of negotiation

➢ Submitting the proposal: proposal for statutes, scientific description, declaration of the host member

    ☐ 2 years of negotiation

➢ ERIC in place until 2040

The EMBRC Story - numbers

- Budget for hosting the EMBRC secretariat is 600K€/ year (host premium) and annual membership fee of 100K€ (for France, paid by ministry).

- The budget agreed upon by our ministry is 2,8M€ in cash for 5 years (through various sources, project etc… not paid directly by the ministry).

- Very strong push and support from French Ministry for research

Liaising with the French Government

- Meeting with the departments in charge of RI within our Ministry of Research

- Satisfied with the ambition of an AISBL or ERIC LE for DiSSCo (not IO)

- More experienced with ERIC than with AISBL

- If the ERIC option is selected, the French National Node will have to submit a file to be evaluated by an inter-ministerial committee (Research, Foreign Affairs, Finance) and representatives of other RIs

- If the ERIC option is selected, the French vote would be assumed by the Ministry. There will be the possibility of consultation prior to the decision making, e.g. people from MNHN could be invited to the GA as experts

- Financing of either model is not guaranteed and is delivered through the institutions in both cases

"Bringing the irreplaceable data stored in natural science collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

The preparatory phase project of DiSSCo Research Infrastructure - Distributed System of Scientific Collections

Funded by the Horizon 2020 Framework Programme of the European Union

https://www.dissco.eu/prepare/

DPP All Hands Meeting 1

The ERIC Forum

Eva M. Alonso.

DiSSCo Prepare Project Manager

Coordination & Support Office

The ERIC Forum

What is an ERIC?

The European Research Infrastructure Consortium (ERIC) is a specific legal form to facilitate the establishment and operation of research infrastructures with European interest. It provides

- a legal capacity recognised in all EU Member States,
- flexibility to adapt to specific requirements of each infrastructure
- a faster process than creating an international organisation, and
- exemptions from VAT and excise duty.

What is the ERIC Forum/Objectives

A H2020 Project aiming to:

- **strengthen coordination** & **networking** reinforcing the informal ERIC network,
- **support the organisation of specific thematic meetings**,
- support ERICs in preparation, based on **best practices**;
- support **common communication and outreach activities** and strengthen external representation of ERICs' as a stakeholder in consultations and other policy actions that could affect them.

Benefits from participation

Composition
- 20 ERICs + 3 pan-European RIs in the process to become an ERIC

Methodology
- Bi-annual meetings to discuss shared challenges in multiple areas)HR, VAT and excise duties, management, socio-economic impact of ERICs, performance monitoring, procurement and more.

Follow and participate in its events may bring advantages:
- knowledge
- save time and resources
- networking for future cooperation
- strength current international relations and extend them
- visibility
- better positioning at national level > increasing trustiness

"Bringing the irreplaceable data stored in natural science collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

The preparatory phase project of DiSSCo Research Infrastructure - Distributed System of Scientific Collections

https://www.dissco.eu/prepare/

Questions ?

Before we have leave this session .... Please join again slido.com and enter the following code event :

https://app.sli.do/event/4orhbkk2

"Bringing the irreplaceable data stored in natural science collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

*The preparatory phase project of DiSSCo
Research Infrastructure - Distributed
System of Scientific Collections*

https://www.dissco.eu/prepare/

TASK 7.3: Develop and establish DiSSCo policies
Wednesday, January 20th 2021

Task lead: Vince Smith (NHM London)

January 20th 2021

AGENDA (http://bit.ly/39N5aJv)

1. Task overview, milestone & deliverable

2. Progress to date (towards a "design blueprint")

3. A vision for the deliverable (policy self assessment tool)*

4. From user stories to definitions, requirements and a design blueprint*

5. Design blueprint (milestone) document sections*

6. Discussion points

*major discussion points

https://docs.google.com/document/d/199dKlQrt4NILLksT5D_SY-J4MIClyczV0__AgtFuTiw/edit?usp=sharing

1. MILESTONES AND DELIVERABLES

▪ **Milestone**

MS7.5 Design of the DiSSCo policy framework tool **(report)**
Due: April 30th 2021 (extended from 29th Jan 2021)

= Design principles and first stage plan for policy framework tool

● **Deliverable**

Deliverable 7.3 - Assessment tools and direction map to the implementation
of common DiSSCo policies **(pilot tool)**
Due: Jul 31st 2022

= Web based tool for institutions to self-assess DiSSCo policy readiness
(i.e. presence, absence and alignment of institutional policies)
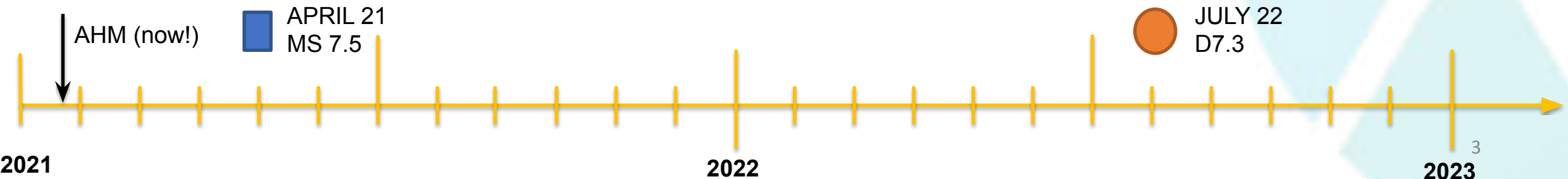
1. MILESTONES AND DELIVERABLES
   - Key reference documents

Past meeting notes - http://bit.ly/38ZMOFW

https://docs.google.com/document/d/13UtDedI7uyYYL2sDtE903iTHRILvGq0kcNUhqfBkY50/edit?usp=sharing

Main planning document - http://bit.ly/3itJCFH

https://docs.google.com/document/d/1qTfZvCoYXcOkIeEJ1mA2Qo7dXqETh6AU8nLJKLJGxXQ/edit

User stories - http://bit.ly/38VyqPe

https://docs.google.com/spreadsheets/d/1A6MMKLoP9rHQVHHuj9CYrjjICFqq5QBkrPPkr2pRXGk/edit?usp=sharing

Policy tool requirements - http://bit.ly/2XV1Oio

https://docs.google.com/spreadsheets/d/1DOEY35KH0o5MGFaPBbtRtnbB9xTF1CnzwRcqkiqBC6I/edit?usp=sharing

1. MILESTONES AND DELIVERABLES
   - Principles / misconceptions

7.3 is not about writing institutional policy

Institutional policies are too complex. We are providing the means to classify, align & id. gaps in institutional policy with DiSSCo services needs

7.3 is not about writing DiSSCo policy

But identifying high level DiSSCo service policy needs is in scope

7.3 the tool must be useful to partners & DiSSCo

We need to incentivise institutions to use the tool by adding value to their policy management & activities

2. PROGRESS TO DATE

✓ Tool "statement of purpose" developed

✓ 29 outline user stories developed & classified

✓ Initial policy classification / metadata scheme (developed in ICEDIG & being refined in SYNTHESYS+ http://bit.ly/3o2TBTz

✓ Exemplar design blueprint documents identified (aiding milestone structure)

3. A VISION FOR THE POLICY TOOL
- Tool "Statement of purpose"

We will develop an online checklist tool which allows a DiSSCo Partner to map their institutional policies against the policy requirements of DiSSCo Services to show policy alignment, and for the DiSSCo CSO to see the overall state of policy compliance and gaps across all DiSSCo Partners.

The tool will:

- support the upload/linkage of institutional policies;
- support deposition of a list of the policy requirements from the DiSSCo CSO for DiSSCo Services;
- contain a classification of terms (metadata schema) for these policies/services;
- provide the self-assessment interface that allows a user to apply the metadata schema and classification terms to their institutional policies, such that they can demonstrate / self certify alignment with the DiSSCo service policy needs.

3. A VISION FOR THE POLICY TOOL
- A possible user journey

DiSSCo Knowledgebase

Institutional Policies

(policies classified/scored via Synth+ metadata schema)

(Institutional metrics)

(services classified/scored via Synth+ metadata schema)

DiSSCo Services

Institution dashboard
(maturity index, gaps, comparisons to peer average, more?)

DiSSCo dashboard
(member institutions, listed, ranked and scored)

Institution process / view

CSO process / view

4. FROM USER STORIES TO A DESIGN BLUEPRINT
- Introducing the policy tool requirements worksheet

Policy tool requirements worksheet
https://docs.google.com/spreadsheets/d/1DOEY35KH0o5MGFaPBbtRtnbB9xTF1CnzwRcqkiqBC6I/edit?usp=sharing

4. FROM USER STORIES TO A DESIGN BLUEPRINT
- Next steps

- Define key concepts and user roles

- Convert user stories (bus. requirements) to functional requirements

- Identify key data requirement needs

- Develop user-access requirements

- Develop reporting requirements

- Identify non-functional requirements

- Writeup into the milestone as the design blueprint

5. DESIGN BLUEPRINT (MILESTONE) DOCUMENT SECTIONS
- Possible document sections

Introduction
    Purpose
    References
Overall Description
    Product Perspective
    Product Functions
    User Classes and Characteristics
External Interface Requirements
    User Interfaces
    Software Interfaces
    Communications Interfaces
System Features
    Data source import and validation
    Interactions with external data services
    Data curation
    Reporting and visualisation
    Analytical extracts
    User administration

Other Non-functional Requirements
    Performance Requirements
    Safety Requirements
    Security Requirements
    Software Quality Attributes
    Business Rules
Other Requirements
    Appendix A: Glossary
    Appendix B: Analysis / Visualisation Models
    Appendix C: Non prioritised requirements

Are there additional sections missing from this list?

6. REVISITING DISCUSSION POINTS
   - For consideration

- Does the vision for the policy tool align with needs / expections?
  For example, this aligns with the need for self assessment tools in WP3?

- Are the tasks (next slide) the right ones & are we missing anything?

- Are the section headers for the milestone correct?
  Are there better blueprint examples we might base this milestone on?

6. DISCUSSION POINTS
  - Assigning these tasks to partners?

- Define key concepts and user roles

- Convert user stories (bus. requirements) to functional requirements

- Identify key data requirement needs

- Develop user-access requirements

- Develop reporting requirements

- Identify non-functional requirements

- Writeup into the milestone as the design blueprint

Can existing partners take these on independently ?

If not, are there alternative useful tasks matching available skillsets?

Instructions for the meeting

- The sessions will be recorded and available after the All Hands meetings;

- By default, all attendants are muted from the start;

- At the beginning of the session, the team must decide who will be the note-taker and chat moderator, if necessary;

- To save the chat, the coveners must download it before the end of the session;

- Link to the general sessions : https://us02web.zoom.us/j/81927609291?pwd=WEpSa0l6VWtqbExYODBSSHJnUWNoQT09

Agenda

Time	Topic	Lead
15.00-15.05	Welcome	AC
15.05-15.20	1) ENGAGEMENT Task 8.1: National Nodes engagement	AC/ MLK
15.20-15.35	2) Strategic engagement	MLK
15.35-16.05	3) D8.2 Thematic Specialisation Plan	MLK/RBINS
16.05-16.20	4) ADVOCACY Task 8.4: Advocacy and Outreach	EA
16.20-16.25	5) Next steps	AC
16.25-16.30	AOB	AC

https://docs.google.com/document/d/1I-OvOLouGCecUqmBw06OVT56gl5yX5Jqc_CB7saAMV0/edit

Engagement

1) **T8.1 National Nodes Engagement**

Task description and role within DiSSCo.

- Identify **contact points**
- Channel their **relevant institutional strategies and policies** to CETAF
- Validate the **resume of the policies corpus** developed under WP7, and ensure its widest dissemination nationally.
- Provide **feedback** for outputs and participate in surveys.
- **Stay informed** on the development of DiSSCo projects

Engagement

1) **T8.1 National Nodes Engagement**

Task Objectives

- **Ensure** alignment and harmonisation
- **Channel** engagement of national governments
- **Gather** the state-of-the-art at national level
  → Foundation for the construction of an overall strategic map for DiSSCo distribution and operation activities.

Engagement

1) **T8.1 National Nodes Engagement**

Achievements: together, we have already...

- **Collected** national priorities and strategies to secure national governments' commitment through
  - priorities matrix, Funders Forum survey & bilateral meetings.
- **Participated** in 9 monthly meetings
- **Launched** internal engagement campaign

Engagement

1) **T8.1 National Nodes Engagement**

Achievements: together, we have also...

- **Developed** External Communications tools as part of T8.2
  - Key messages webpage
  - DiSSCo Brochure
  - DiSSCo powerpoint presentation template

(in depth presentation during NNs meeting in Feb'21)
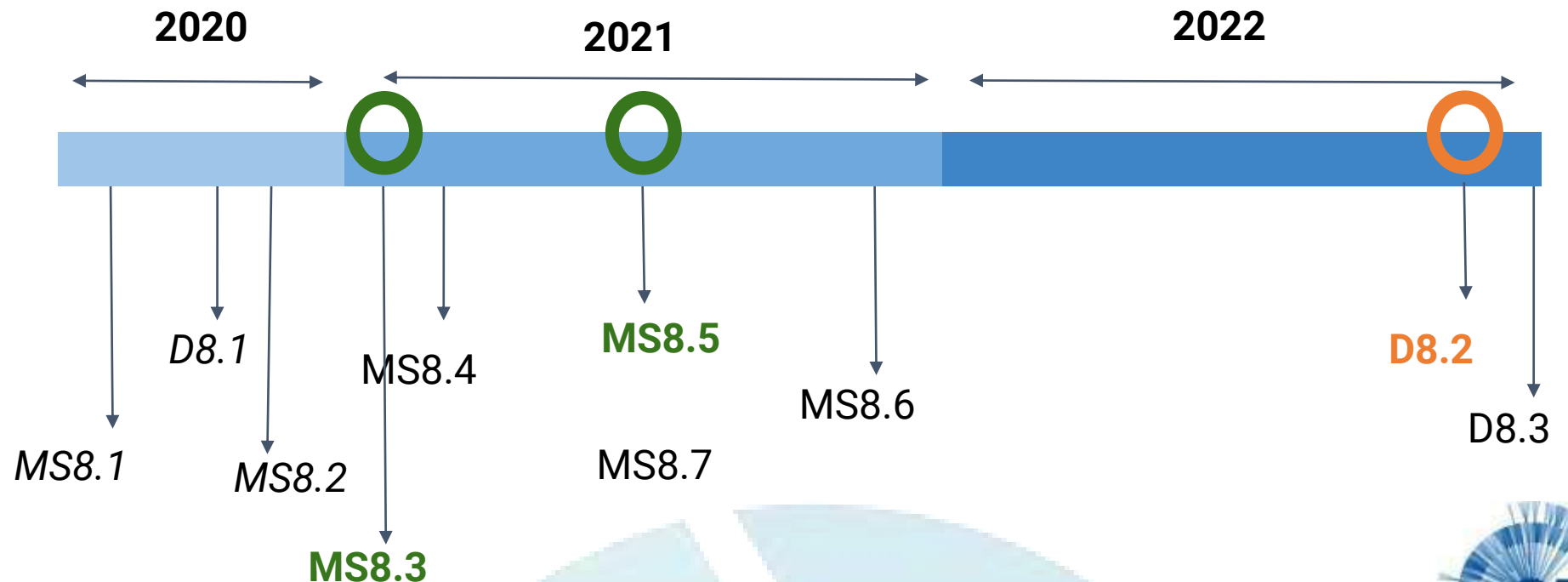
Engagement

1) **T8.1 National Nodes Engagement**

Challenges: learning opportunities…

- Involving the institutional **communication** departments
- Revise **mechanisms of engagement** of NNs during and beyond monthly meetings
- Clarity of the **scope of DiSSCo** related projects
  - better understanding of the landscape of the preparatory phase

Engagement

1) **T8.1 National Nodes Engagement**

Timeline of outcomes (T8.1) **and others in WP8**

2020   2021   2022

D8.1   MS8.5   D8.2

MS8.4   MS8.6   D8.3

MS8.1   MS8.2   MS8.7

MS8.3

WP8 Connected Milestones & Deliverables

- **MS8.1** Website Operational (T8.2) *- submitted*

- **D8.1** Communication Strategy (T8.2) - *submitted*

- **MS8.2** First communication strategy campaign launched targeting national nodes engagement (T8.2) - *submitted*

- **MS8.4** Setting criteria for procurement framework (T8.3) - *due April '21*

MS8.3 National & Institutional level strategies collected - *due Feb '21*

MS8.5 Initial findings for specialisation plan - *due July '21*

- **MS8.6** Identifying indicators for alignment (T8.3) - *due Nov '21*

- **MS8.7** Outreach and Advocacy Strategic plan v1 in place (T8.2)- *July '21.*

D8.2 Thematic Specialisation Plan - *due Sept '22*

- **D8.3** Partnership Best Practices (T8.3) - *due Nov '22*

Engagement

2) **Strategic Engagement**

- Current SE mechanisms: meetings, surveys, matrices
  - lessons learned.
- Why implement new SE mechanisms?
  - we want to hear more from you
  - Way forward:
    - newsletter: DiSSCo updates/ DiSSCo explained
    - other

Engagement

3) D8.2 Thematic Specialisation Plan

Task leading partners: CETAF, RBINS, LUOMUS, Naturalis.

A) Shared definition

"The organizations contributing to DiSSCo form a very rich but diverse network. Consequently, many assets are unevenly distributed amongst these organizations. In order to document, promote and optimize the assets, and possibly to identify gaps, we will develop a tool to collect all relevant information and enable the assessment of the institutions specificities.

The specialization plan will come out from the information in the tool, the gaps identified or SWOT analyses made."

Engagement

3) D8.2 Thematic Specialisation Plan

B) Information collation (1)

- type and size **of collections**

- digitisation techniques and capacity

- training (capacity and domains)

- strategic goals of the institutions (services, policies and willingness to contribute)

- ...

Engagement

3) D8.2 Thematic Specialisation Plan

B) Information collation (2)

- to align with the definition and needs of the Specialisation Plan

- to identify which part of information collected from NNs (MS8.3: National priorities and strategies) might be of use and application for D8.2 (RI distributed organization and operation).

- to link to the Policies metadata schema (SYNTHESYS+ T2.1)

- Launch to **capture and refine** available information (institutions)

- **Sources:** other WPs (WP8 NNs priorities, WP2 Training, WP7 Policies), among others..

Engagement

3) D8.2 Thematic Specialisation Plan

C) Development of an assessment tool

*Connected to **MS8.5** Initial findings for the specialisation plan - due Aug 2021.*

- to define criteria and requirement

- **to ensure** technical feasibility

- **to develop** mechanism

- **to test**

This tool will help to retrieve and combine data but also enable institutions to perform self-assessment on their specialisation assets.

Engagement

3) D8.2 Thematic Specialisation Plan

D) Analysis of the tool potential uses

- defining use cases

- implementing high-level combination of information

E) Thematic Specialisation plan

Advocacy Strategy

4) T8.4 Dissemination, Outreach & Advocacy

Task description (*advocacy*)

➢ Develop and implement an strategic plan to address national authorities.

Objective (*advocacy*)

➢ Liase effectively with national authorities to ensure their practical commitment & alignment with national priorities

  ➢ **The Funders Forum Advisory Board (FF)**

Advocacy Strategy

4) T8.4 Dissemination, Outreach & Advocacy

Key Actions

- ➢ Collect and analyse information provide by NNs (NNs consultations)
- ➢ Advocacy Strategy > Phase I – July 2020
  - ➢ **FF information package** (online/hard copies) & support letters
  - ➢ Tailor-made national advocacy actions
    - ➢ **Bilateral meetings (>21)**,
    - ➢ **Peer to peer advocacy** (NWO – The Netherlands)
    - ➢ Country fact-sheets
- ➢ Supported by communication tools (key messages, social media, brochure, etc.)

Advocacy Strategy

4) T8.4 Dissemination, Outreach & Advocacy
Bilateral meetings

➢ Better understanding of the national priorities - national fact-sheets,
➢ Common development of national advocacy actions - country specific,
➢ Trigger regional discussions (Scandinavian countries) // discussions on DiSSCo position

Peer to peer advocacy actions
➢ To facilitate national engagement (UK case)
➢ Led by the Dutch Research Council (NOW)

Advocacy Strategy

4) T8.4 Dissemination, Outreach & Advocacy

The Funders Forum shall:

- **inform** on matters related to legal and financial recommendations of DiSSCo RI
- **enable** DiSSCo to effectively adjust its development to national and international priorities
- **facilitate** the key stakeholders endorsement of the core implementation and operational principles
- **enable** the national representatives to make informed decisions on their future financial commitment to DiSSCo and its relationship with the scientific institutions
- **enable** a more consolidated degree of engagement and cooperation between DiSSCo and the scientific institutions

Advocacy Strategy

4) T8.4 Dissemination, Outreach & Advocacy

Key achievements – Phase 1

- **DiSSCo awareness** in 21 countries
- **Funders Forum Advisory Body inaugural meeting**
  - Online meeting on February 25th.
  - Current representation (8 countries) - **BELGIUM, GREECE, U.K, SLOVAKIA, DENMARK, BULGARIA, ESTONIA, THE NETHERLANDS**
  - +3: **FRANCE, ITALY, LUXEMBOURG**
  - In discussions: **PORTUGAL, FINLAND, SPAIN**
- Together we are building an even stronger community based on a common understanding

5) Next steps

- NNs monthly meetings
  - Feb'21: Presentation of Communication material
  - March'21: Review of FF meeting
- WP outcomes:
  - MS8.3-Strategies collected (Feb'21),
  - MS8.5-Specialisation Plan initial findings (Jul'21)
- Launch of **T8.3** Stakeholders Engagement
- Towards **iGA3** and WGs (Strategic positioning and Conc
- Close linkage with **WP7** (LE and Governance)

6) AOB

"Bringing the irreplaceable data stored in natural science collections to life and enabling research at an unprecedented scale"

THANK YOU FOR YOUR ATTENTION !

Logo of the Institution

The preparatory phase project of DiSSCo Research Infrastructure - Distributed System of Scientific Collections

https://www.dissco.eu/prepare/

Task 8.3 DiSSCo stakeholder engagement

Speakers: Laurence Livermore & Vince Smith (NHMUK)
Partners: CETAF, Luomus, MeiseBG, MNHN
Note taking: Helen Hardy (NHMUK)

Funded by the Horizon 2020
Framework Programme
of the European Union

Session Aims

- Task overview

- Discuss planning & scoping document

- Discuss subtasks & subtask leadership

- Agree next steps

"Task 8.3 will strengthen the linkages and build strategic partnership frameworks with the relevant external stakeholders and communities."

Broad scope including:
- Review procurement strategy for DiSSCo external collaboration (e.g. MoUs) and tendering
- Identification of synergies, coordinated actions, and joint service development
- Creating framework for alignment and cooperation and a set of best practices for partnerships
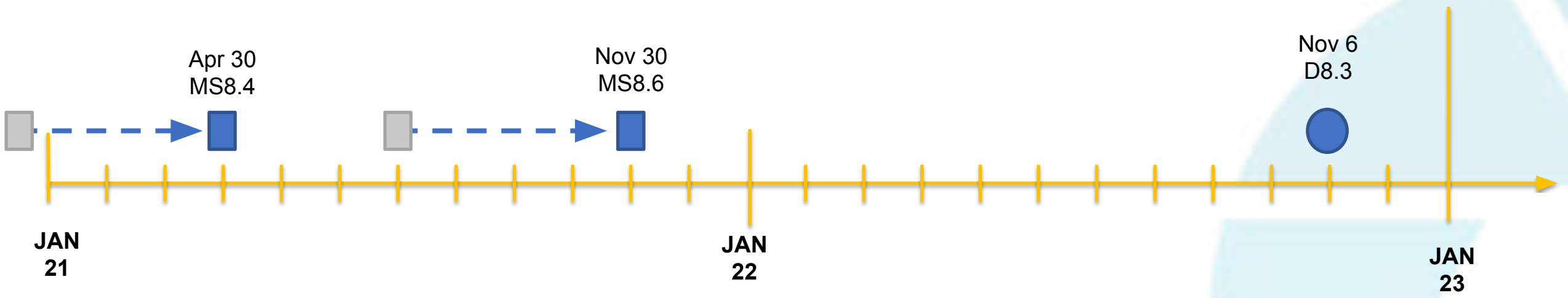
Overview of DiSSCo General Assembly task force on Strategic Positioning
(Vince Smith)

Proposed subtasks

- **8.3.1 Procurement** – Strategy and Policy = MS8.4 (April 2021)

- **8.2.2 Stakeholder Analysis**

- **8.2.3 Stakeholder Engagement**

Task Member	Person Months
NHM	4.5
CETAF	1.5
Luomus	1
MeiseBG	1
MNHN	1

Next steps

- Agree Task Team meeting frequency – every four weeks?

- Next three months focus on T8.3.1/MS8.4

- Create/update/review timeline of strategic DiSSCo (financial) requirements

- Review and integrate strategic partnerships from General Assembly work

- Collate best practices