H2020-INFRADEV-2019-2
Grant Agreement No 871043

Corpus of previous studies on socioeconomic impact compiled

DiSSCo Prepare WP1 – Milestone 1.4

Authors

Rui Figueira (ULisboa), Elsa Fontainha (ULisboa), Sofie De Smedt (MeiseBG), Ana Casino (CETAF), Patricia Mergen (MeiseBG), Elspeth Haston (RBGE)

Contributors

Henrik Enghoff (NHMD), Alexandra Marçal Correia (ULisboa)

ABSTRACT

This Milestone 1.4 report for DiSSCo Prepare Work Package 1 Task 1.4 provides the review of recent frameworks and studies of socio-economic impact of research infrastructures. It also adds a compilation of socio-economic impact indicators recommended or used for the assessment of research infrastructures. The report includes the review of the analysis of impact assessments of research infrastructures and institutions analogous to the goals and domain of activity of DiSSCo. It also includes a definition of the scope of DiSSCo, the areas of impact, user communities and services, which help to identify the relevance of indicators to be selected. It finally provided a list of the actions to be developed in Task 1.4 towards the identification of a significant list of recommended socio-economic impact indicators to be used by DiSSCo.

KEYWORDS

SOCIO-ECONOMIC IMPACT,

RESEARCH INFRASTRUCTURES,

INDICATOR COMPILATION,

KEY PERFORMANCE INDICATORS

INDEX

01 INTRODUCTION

This milestone serves the purpose of reviewing, compiling and aggregating the existing bibliography, frameworks and indicators of the socio-economic impact (SEI) of research infrastructures (RI). The main goal of this compilation is to support the definition of a set of indicators to be implemented for the SEI assessment of DiSSCo.

The need to perform a credible socio-economic impact assessment is justified by the demand for understanding and evaluating the return on investment in these facilities, to support informed decision-making and RI management with useful information for negotiations with funders (OECD, 2019). In the case of RI that are part of ESFRI Roadmaps, this is further a requirement in their regular assessment of the scientific case (ESFRI, 2021).

For RI, the scientific output is the most important, but the SEI has a broader scope, by including cultural, educational, economic and social impacts. However, SEI of different RI should never be compared because of their uniqueness (Hajdinjak, 2019). The SEI assessment faces several challenges that need to be considered (Hajdinjak, 2019; OECD 2019):

- difficult to perform in cutting edge fields
- RI targets multiple stakeholders
- research outcomes uncertain and non-linear
- time lag between research and impact
- difficult to gather data about impacts and to verify them
- impacts can be direct and indirect, intended and unintended
- changes during the lifecycle of the RI
- the relevant types of impact varies depending on RI specific goals
- societal impact may be broad and difficult to measure and allocate monetary value.

Furthermore, there might be legal barriers related to non profit status, limitations of commercial activities and profit making depending on the legal forms of the ERIC and the member institutions of the RI.

There are several approaches to measure SEI, but no single method can appropriately meet the information needs for that assessment (Vignetti et al., 2019). Furthermore, many RI implement monitoring schemas using Key Performance Indicators (KPIs), but fewer collect indicators for impact assessments (Vignetti, 2021). Although connected, impact assessments are not identical to monitoring. The performance monitoring is a continuous process generating data to track the progress of an action, while the impact assessment is a structured process that takes place at a given point in time, allowing to assess the implications (past, future or both) of proposed actions (Vignetti, 2021).

DiSSCo crosses two domains, environment (biodiversity and geodiversity) and digital data, in the scope of museums, particularly natural history collections. Therefore, impacts could result from both domains, possibly signergeticly. A SEI study targeted to assess, by a cost-benefit analysis, the benefits of improving the knowledge about biodiversity in Australia (Deloitte Access Economics, 2020), namely the discovery and documentation of species, indicate that for each dollar spent, benefits range from 4 to 35 dollars. The return results from an increase in biosecurity diagnostics, e.g. reducing the frequency of genuine threats, from biodiscovery for human health, agriculture R&D and biodiversity conservation.

These results can be leveraged by the digitisation of collections, which increases accessibility and usability of data for knowledge production. In another study (Popov, 2021), it was found that the full digitisation of the London's Natural History Museum collections would give a return of seven to ten times on the investment, with a benefit of 2 billion pounds over 30 years. This figure resulted from the analysis of the impact of digitisation on five areas: biodiversity conservation, medicines discovery, invasive species, agriculture R&D and mineral exploration.

In addition to these, it is reasonable to expect additional increased impacts where natural history collections already have an important role, like on defining baselines and time-series for key environmental variables, on training and education, and on scientific communication. Several reviews have discussed the use of natural history collections with examples (Brooke 2000; Suarez and Tsutsui 2004; Tewksbury et al. 2014; Rocha et al. 2014).

DiSSCo aims to digitally unify all European natural science assets under common access, curation, policies and practices. It will create a unique access point for integrated data analysis and interpretation through a wide array of digital services provided by its community. This sharing and harmonisation of practices and processes will promote capacity transfer between countries and between bigger and smaller institutions, contributing to level up their capacity and the accessibility of their collections.

DiSSCo will enable the data to be easily Findable, Accessible, Interoperable and Reusable (FAIR principles). For researchers, this will make access to data instantaneous, and in many cases reduce costs and impacts of travelling. But the digital transformation of museums can also occur in its interaction with the public, which, through the exploration of the Digital Specimen concept and digital technologies, can make passive visitors into active participants. Ruttkay and Bényei (2018) provide examples how digital technologies in museums can promote motivation and engagement, education in different ways, learning by doing, participation, adaptation to different visitors and extension in time and space. Digitisation not only enables access to objects in different ways, but also enriches it with metadata.

The digitization and FAIRification (see glossary) of Natural History Collections will create large volumes and variety of data (*big data*). This will support existing scientific and knowledge creation activities, in environment, biodiversity, and related domains mentioned before. It is, additionally, a possible pool for new data-driven innovation (OECD, 2015) resulting from machine learning and artificial intelligence applications. DiSSCo RI may well turnout to be a data infrastructure in the sense defined by OECD (2015), which potential for value creation are based on the following properties of data: *i) the (non)rivalrous nature of their consumption, ii) their (non)excludability, and iii) the economics of scale and scope in the creation and use of data*. In fact, data aggregation and access through DiSSCo can be seen as an infrastructure resource, meaning a non depletable capital good and with a theoretical unlimited range of purposes, even outside the domains of its origin. It is necessary, nevertheless, that data is under a governance framework for better data access, sharing and interoperability, a component of the DiSSCo RI construction and implementation.

The SEI of all of these dimensions should be captured by indicators. This report aims to be a step towards the definition of a framework of indicators to be used in the DiSSCo SEI. The following sections of the report include i) the methodology used to make the compilation of SEI indicators, ii) the review of existing frameworks of SEI for RI, iii) the review of relevant SEI studies applicable to the domain of biodiversity and natural history collections, iv) the identification of the areas of impact, users and services of DiSSCo, v) the result of the compilation, and vi) the identification of next steps in the definition of the DiSSCo SEI indicators.

02 METHODOLOGY

The compilation of SEI indicators followed the workflow described in Figure 1.

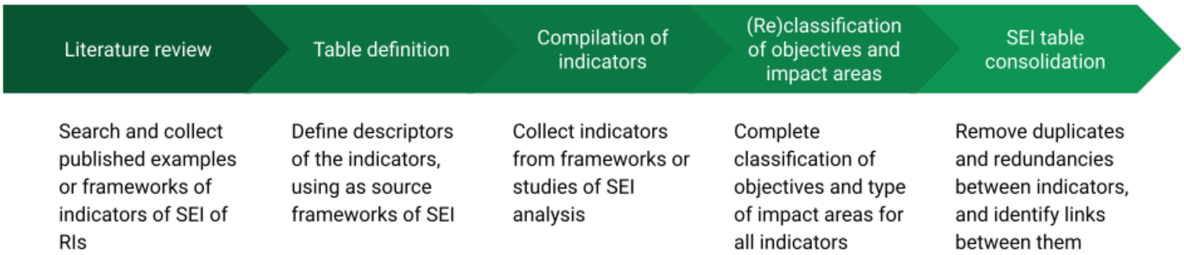

Figure 1. Workflow of the compilation of the SEI indicators.

The first step for the compilation of SEI indicators was to identify existing frameworks specifically developed for the impact assessment of RI, in particular, those applied to ESFRI projects or landmarks. In addition to these, a more extensive bibliographic review on SEI exercises practised in other types of organisations or institutions, and initiatives, was performed, particularly in the domain of environment, including biodiversity.

The review enabled the preparation of a template for the compilation of the parameters related to the indicator description. The initial source of inspiration for that template was the ESFRI framework (ESFRI, 2019), which was later expanded to accommodate other parameters of socio-economic impact categories used by other frameworks (OECD, 2019; Helman et al., 2020, Alluvium, 2016), or additional operational terms. The final template includes 37 columns described in Table 1.

The compilation of indicators was done by transcription of the original source, without reinterpretation or edition of the text. However, in the cases where the indicator was presented in an aggregated form, as in Alluvium (2016), it was necessary to desegregate it, so that it became equivalent in scope to the indicators of other frameworks. The details of information in the description of indicators were different between the different frameworks, with the ESFRI document being the most detailed. At this stage, only the objectives and impact areas of the different frameworks were interpreted in order to classify all indicators compiled, because this is useful to identify and select the most relevant indicators for DiSSCo.

Finally, the compiled table was duplicated to create, through a consolidation step, a cleaner table that removes duplications of indicators between frameworks, as well as redundancies between similar indicators. Nevertheless, these duplications were documented in an auxiliary table for future reference. In the consolidated table, the columns of the original template that were not filled in due to the lack of information in the sources were removed. The consolidated table retains 24 columns for the description of the indicators.

The full file of compiled indicators is available in Appendix 1, and also as a web data source (https://tinyurl.com/DISSCO-SEIcompilation), and contains the following sheets:

- **metadata**: description of the tables and columns included in the file;
- **list_indicators**: List of all indicators found in the four frameworks consulted;
- **list_consolidated**: Consolidated list of indicators, obtained after removing duplications and redundancies between indicators of different frameworks. Only columns with meaningful information at this stage are kept in this consolidation. The columns of this table have the same meaning as in list_indicators;
- **related_indicators**: links between duplicated or redundant indicators of different frameworks;
- **references**: references list of the consulted frameworks.

Table 1. Template table for the compilation of the SEI indicators. Also available as a web data resource at https://tinyurl.com/DISSCO-SEIcompilation.

Column_name	Column_long_name	Description	Reference	Example
ID_list_indicators	ID list of indicators	ID primary key for table list_indicators		
ID_related_list_indicators	ID related in the list of indicators	ID of the related indicator, in an alternative framework. This is a foreign key to ID_list_indicators		
ID_source	ID of the source	unique ID, with indication of the source of the indicator. If the source has IDs for the indicators, these are used. A prefix of the source framework is added.		
ID_original	ID original	ID of the indicator at the source		
Indicator_Name	Indicator Name	name of the indicator	ESFRI (2019)	Example: “3D’s images delivered to educational purposes”
Indicator_code	Indicator code	code of the indicator		Example: A. m.3D (Activity of delivering images 3D)
Indicator_short_name	Indicator short name	acronym of the indicator		Example: i3D
Related_Indicator_source_ID	Related Indicator source ID	ID_source of the indicators that are similar or related to the current indicator		
Type_of_indicator	Type of indicator	type of indicator, according to Ri-PATHS framework. Assumes one of the following values: activity, outcome, impact	https://ri-paths-tool.eu/en/glossary, Helman et al. (2020)	
Definition	Definition	definition of the indicator	ESFRI (2019)	Example: This indicator measures the images 3D provided by the collection-holding institutions for educational purposes

Column_name	Column_long_name	Description	Reference	Example
Rationale	Rationale	reason or objective for using the indicator	ESFRI (2019)	Example: This indicator describes one activity of disseminations of the collections, identifying the users and quantifying the volume of that activity. It can be used for.... The main limitations are : ...
Objective	Objective	identifies the objective of the RI that this indicator helps to monitor. This field is based on the ESFRI framework, and it was interpreted for the indicators of the other frameworks. Uses the following categories: Enabling Scientific Excellence, Delivery of education and training, Enhancing transnational collaboration in Europe, Facilitating economic activities, Outreach to the public, Optimising data use, Provision of scientific advice, Facilitating international cooperation, Optimising management, Enhancing Collaboration in Europe	ESFRI (2019)	
Impact_area	Impact area	area of impact, according to Ri-PATHS framework. Assumes one of the following values: Human Resources, Economy and Innovation, Society, Policy	https://ri-paths-tool.eu/en/glossary, Helman et al. (2020)	
Category_of_SEImpact	Category of Socio-Economic Impact	category of socio-economic impact measured by the indicator, according to the OECD framework. Uses a controlled vocabulary: scientific, training and education, economic, social and societal, technological	OECD (2019)	
Nature_of_indicator	Nature of the indicator	nature of the indicator: numeric, binary, categoric, narrative		
Scope_of_indicator	Scope of the indicator	acceptable range of values for the indicator		min, max
Temporal_scope_of_	Temporal scope of	temporal scope of the impact of the indicator		long run; short run

Column_name	Column_long_name	Description	Reference	Example
indicator	the indicator			
Assumptions	Assumptions	assumptions or limitations that need to be considered when using the indicator	ESFRI (2019)	
Unit_of_measure	Unit of measure	unit of measure of the indicator. Example: euros, number of items	ESFRI (2019)	example: numeric
Relative_Measure_Associated	Relative Measure Associated	relative measure of the indicator		examples: 3D's images delivered to educational purposes per user with educational purposes
Data_Information_needs_and_Resources	Data/Information needs and Resources	data requirements to measure the indicator	ESFRI (2019)	example: information about the requests by user (for educational purposes)
Who_provides_information	Who provides the information	name of the source of the indicator	ESFRI (2019)	Examples: National Statistical Office; the RI
Method_of_gathering_data	Method of gathering of data	method used to obtain the indicator		Example: Questions/Survey associated (type of answer Y/N or other)
Logistic_requirements	Logistic requirements	logistic requirements to obtain the data		Example: Registration (digital request form online? Papers request form?: fill by user? Fill by services?)
Indicator_Calculation	Indicator Calculation (detailed methodology for)	mode of calculation of the indicator		Example formula and meaning
Disaggregated_data	Disaggregated data (broken down by detailed sub categories)			Example1: Botanical Gardens, Research Institutes, Museums, etc.; Example 2: size categories of RI

Column_name	Column_long_name	Description	Reference	Example
Estimated_Cost_Data_Collection	Estimated Cost of Data Collection	direct costs to collect information to obtain the indicator	ESFRI (2019)	Example costs for RI: zero; Low; Medium;High ; NOTE: where are the direct returns of the services provided registered?
Level_reporting_burden	Level of reporting burden	indication of the effort needed to report the indicator	ESFRI (2019)	
Frequency_of_measurement	Frequency of measurement		ESFRI (2019)	Example: bi-annual
Spatial_scope	Spatial scope	spatial scope for the calculation of the indicator		Example: for the main RI and collection regional delegation
Last_updated	Last updated	last update of the indicator definition		Example: 01.03.2021
Internal_users_of_indicator	Internal users of the indicator	internal users of this indicator		Example: IU:123XPTO
External_users_of_indicator	External users of the indicator	external users of the indicator		Example:EXU:456XPTO
Assessment_of_Indicator_quality_and_comparability	Assessment of Indicator quality and comparability	indicates whether the indicator is of common use, or needs to be refined to be applied	ESFRI (2019)	
ID_Citation	ID Citation	citation ID of the bibliographic source for indicator use and calculation, listed in references table		
Quality_Control_of_Indicator	Quality Control of the Indicator	method of QC/QA of the indicator		Example: How? Who? When?/Frequency
Additional_Issues_Observations	Additional Issues or Observations		ESFRI (2019)	

03 SOCIO-ECONOMIC IMPACT FRAMEWORKS

Three frameworks on the socio-economic impact assessment of RI were recently published (ESFRI, 2019; OECD, 2019; Helman et al., 2020). In addition to an extended review, these frameworks resulted also from consultations with RI through surveys, workshops or pilot studies. These prompted a useful and comprehensive list of methods and indicators that serve as a basis for the DiSSCo SEI development. We provide here a summary of the three frameworks from which indicators were compiled for this milestone. In addition to these frameworks, some other frameworks (Hajdinjak, 2019) or SEI exercises performed by RI (e.g. ACTRIS) were also consulted, not adding, however, additional different indicators to the list.

3.1. ESFRI RI performance monitoring

In an attempt to develop a common approach to measure the performance of RI, ESFRI adopted a Working Group report published in 2019 (ESFRI, 2019) that proposes 21 key performance indicators (KPI). These can be voluntarily adopted by a RI, and although being formulated as KPI to measure performance, they can also be used as a proxy to measure impact. The adoption can be decided if they are aligned or adapted to the objectives and activities of the infrastructure, and follow the RACER criteria (European Commission, 2015):

- Relevant – i.e. closely linked to the objectives of the RI over a particular period of time.
- Accepted by the RI (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

The list of KPIs proposed are organised by nine objectives that reflect several aspects relevant to RI. This relevance varies depending on the type and scope of the RI and phase of its life cycle. It is worth noting that the ESFRI WG proposal is targeted to the assessment of operational (landmark) RI. Its application to early phases implies adaptation. Table 2 includes a list of those objectives and their relevance to DiSSCo.

Table 2. Objectives of an RI identified by the ESFRI report and their preliminary relevance to DiSSCo.

Objective	Relevance	Phase of life cycle
Enabling scientific excellence	high	all
Delivery of education and training	high	all
Enhancing transnational collaboration in Europe	high	all
Facilitating economic activity	medium	operation
Outreach to the public	high	all
Optimising data use	high	all
Provision of scientific advice	medium	construction, operation
Facilitating International co-operation	high	all
Optimising management	medium	preparation, construction

The set of 21 KPIs proposed are all quantitative, e.g., number of users served, number of publications, number of publicly available datasets, to name a few. For each indicator, a data sheet details the relevant information about its definition and description (Table 3).

Table 3. Attributes defined for each ESFRI indicator datasheet.

Indicator (name)
Definition(s)
Rationale
Assumptions
Data/information needs and resources
Who is providing this information
Detailed methodology for indicator calculation
Unit of measure
Frequency of measurement
Assessment of indicator quality and comparability
Estimated cost of data collection (including access to external databases)
Level of reporting burden
Additional issues or Observations

The structure of the datasheet is useful and will be adopted for the description of the indicators to be used by DiSSCo. Nevertheless, the report includes in the Annex 4 other possible indicators by objective, which are of narrative or boolean type. Some of these can be more appropriate when valuing impact.

3.2. OECD reference framework

The reference framework was developed in the scope of the OECD organisation (OECD, 2019), by an Expert Group of the Global Science Forum. It aims to provide funders, decision-makers and RI managers with a tool to evaluate the achievement of scientific and socio-economic objectives, in order to facilitate communication and reporting between RI stakeholders. The tool proposes a set of 25 core impact indicators that can be adopted regardless of the activity area of the RI or its life cycle phase. The list is drawn from a larger set of 58 standard indicators referred by the RI surveyed in support of the study.

The analysis provided by the report identifies the different interested stakeholders and their main interest in an RI impact assessment exercise. These stakeholders and interest include:

- funders (national and/or regional authorities, other funders) - justify the investment and value for money
- implementers (creators, managers or hosts of the RI) - demonstrate value of the RI and impact
- scientific community - foster scientific knowledge
- civil society - value for money and new scientific knowledge

In alignment to the ESFRI framework, the OECD report also identifies several strategic goals of the RI, not limited to the scientific output, but which includes cultural, educational, economic and social impact, revealing the broader scope of the RI. However, the practice of the SEI assessment needs to be connected to strategic objectives: useful, reliable, meaningful, practical and recognised (for economic indicators).

The proposed framework is based on a logic model, adapted from the theory of change, set as a pathway that includes the following steps:

- Inputs: the resources mobilised by the RI to perform its activities. Resources may come from multiple sources and in-kind support can be an important input.
- Activities: what RI do - supporting science and technology, targeting economic and social activities and developing the skills and competencies of human resources.
- Outputs: the results of RI activities: scientific, educational, collaborative and economic.
- Impacts: intended and unintended effects of the RI' activities and outputs over their lifecycle. Activities and outputs can lead to long terms impacts on different aspects of society and the economy

The reference framework includes 25 Core Impact Indicators (CII), which with the additional 33 indicators comprise 58 standard indicators. The former provide a general picture of the SEI of the RI at a certain time, while the full set of 58 indicators reflects the diversity of indicators used regularly by the infrastructures surveyed for the study. The CII are organised around the following sets of impact and strategic objectives categories:

Impact categories

- Scientific impact
- Technological impact
- Economic impact
- Training and education impact
- Social and societal impact

Strategic objectives

- Be a national or world scientific leading RI and an enabling facility to support science
- Be an enabling facility to support innovation
- Become integrated in a regional cluster/in regional strategies / be a hub to facilitate regional collaborations
- Promote education outreach and knowledge transfer
- Provide scientific support to public policies
- Provide high quality scientific data and associated services
- Assume social responsibility towards society

The report provides, for each indicator, the category of impact and strategic objective it belongs to, a detailed explanation of the indicator and the data needed with possible sources of information.

3.3. RI-PATHS Impact assessment framework

The Ri-Paths framework was developed in the scope of the European project with the same acronym (Helman et al., 2020). The framework proposes the impact assessment around several components. The first and most important are the impact pathways, which can be defined as *simplified causal chains of events that connect the activities carried out on a Research Infrastructure to identifiable effects on the economy and wider society*. Thirteen impact pathways are identified in the framework (Table 4), distributed around three main strategic objectives, namely, enabling science, problem solving and science and society.

Table 4. Pathways defined in the RI-PATHS framework to enable RI impact assessments.

Enabling science
Publication-citation-recognition
Employment, operations & standardised procurement
Technology transfer and licensing
Learning and training through joint development of instruments and tools
Learning and training by using RI facilities and services
Training and higher education cooperation
Problem-solving

Interactive problem-solving for the private sector (industry)
Addressing societal and public-sector challenges
Provision of specifically curated/edited data
Science and society
Changing fundamentals of research practice
Creating and shaping scientific networks and communities
Promoting engagement between science, society and policy
Communication and outreach

Not all pathways apply to all RI, the identification of the appropriate ones should be done by the RI, in accordance with its mission and type of RI - virtual or physical facilities, single-site or distributed. The Ri-PATHS project developed an online toolkit, available at https://ri-paths-tool.eu, to guide RI on developing their assessment exercise. In the tool, the details of each pathway indicate relevant stakeholders and a long and comprehensive list of indicators that can be considered for the specific path, which are arranged in four impact areas (Table 5).

Table 5. Impact areas considered in the RI-PATHS toolkit.

Impact area	Dimensions considered	Number of indicators
Human Resources	Research jobs and career development; Skills development for non-scientific staff and users; Relationship capital and international collaboration; Better working conditions; Wider effects	31
Economy and Innovation	Business and industry; Labour market and productivity; Technology transfer and innovation; Impact on the local and regional economy	36
Society	New solutions, technologies, open access data and software for societal use; Knowledge benefits for society in different domains; Public awareness and engagement; Cultural impact; Social inclusion; Environmental impact	17
Policy	Policy, regulations, standards and institutions; Science diplomacy; Co-funding and sustainability; Ethics and trust in science	18

The various indicators, which are quantitative or qualitative, measure (or are a proxy for) different levels of the RI impact. An indicator can measure an activity, an outcome or an impact, which, in the scope of this framework, are defined as:

Activity – Initiatives and endeavours undertaken using the resources of a Research Infrastructure or work performed by Research Infrastructure staff.

Activity indicator - Indicators that capture the scale and nature of a Research Infrastructure's activities; a measure that should form part of internal reporting. The indicators of this type can be considered KPIs, as in the case of the ESFRI framework.

Impact – Intended and unintended long-term effects of activities using the resources of a Research Infrastructure or work performed by Research Infrastructure staff.

Impact indicator - An indicator that reflects the extent and nature of generated effects in the economy and wider society; with few exceptions, impact indicators are estimations.

Outcome – Longer-term effects that stem from the stakeholder uptake of or interaction with Research Infrastructure outputs.

Outcome indicator - Indicators that document the result of the first productive interactions; collecting data by reaching out to involved stakeholders, e.g. via a survey, interview, external reporting or other data-gathering means.

The framework also provides a list of possible sources of information to support indicator calculation. This includes internal or external tracking of several parameters related to staff, users, visitors, costs, publications, citations, appearance in media and social media, events, etc, or performing surveys. The approaches for data analysis include assessment based on impact multipliers, cost-benefit analysis (CBA), approaches based on multiple criteria, theory-based approaches, case studies and narratives, input-output models and methodologies grounded in the knowledge production-function approach. These methodologies, if they are to be adopted to evaluate the impacts, also need specific information to be applied. For example, the data collection must include for monetary information for costs and benefits in the case of CBA or ways to convert information into monetary units.

Cost-benefit analysis (CBA) is used to test whether a project or policy is socially profitable. It is often used to compare alternatives when it is expected that the projects or policies have social impact. To apply the methodology the social costs and the social benefits need to be quantified, usually in monetary units. The costs and the benefits can be tradable or not in the market and, consequently, there are direct pecuniary costs and benefits and also non-pecuniary effects. These non pecuniary costs and benefits are also referred to as negative and positive externalities of the project. Examples of non market costs of projects (promoted by private or public entities) are environmental impacts like pollution or disturbance of wildlife. In the case of digitising natural history collections, examples of non market benefits are preservation of the physical specimens in archive (lower frequency of handling) and reductions in travelling time for the researchers. The CBA is frequently applied to support decisions about subsidising projects with expected social impact. When the social benefits exceed the social costs this can justify the attribution of a public subsidy if the project is not privately profitable or even when it is privately profitable.

Multi-criteria analysis (MCA) is applied to select alternatives adopting a set of different criteria each with a weight. This MCA contrasts with CBA because the objectives are not aggregated in a single objective. The MCA considers *m* alternatives to be assessed based on *n* attributes. One possible way to implement the MCA is (European Commission 2008, p.66; Johansson & Kristrom, 2016, pp. 202-204; Greco, Ehrgott, & Figueira, 2016): i) quantified objectives are defined (not redundant but could be alternatives); ii) to each objective a weight is allocated (for example the relative importance given by research policy); iii) definition of an appraisal criteria (e.g. based priorities by the stakeholders); iv) impact analysis it means that for each criteria (e.g. environmental protection) is indicated the effect; v) forecast of the effects of the policy on each criteria allocating a score; vi) for each stakeholders group is evaluated the associated preference function (it means, the weights) for each criteria; vii) The project (or policy) impact is aggregated based on the sum (or other method non-linear. The following table illustrates the methodology in a case of digitalization of the collections of a given museum.

Criterion*	Score**	Weight	Impact
Biodiversity Conservation	2	0.6	1.2
Medicines Discovery	1	0.2	0.2
Improve Mineral Exploration	4	0.2	0.8
Total		1.0	2.2

* Criteria associated to the key areas, for example. Other objectives: Equity in the access to collections; improve publication, etc.

** Score: 0=none; 1=Scarce; 2=Moderate; 3=Hight; 4; Very Hight.

The project value aggregated is 2.2. And this can be compared with another project using the same approach. Another project with more than 2.2. Will be preferred to this one. However, as this very simple example shows, the results and the selection are very sensitive to the ranges of the score (in this case 0-4) and the weights values.

04 OTHER SEI ANALYSIS RELEVANT TO DISSCO

Some of the assessment exercises of RI or organisations related to or associated with the activity of DiSSCo might provide a good example of the approach and type of indicators relevant to determine the infrastructure SEI. In this particular case, the impact of digitisation and data infrastructure is particularly adequate, as are the cases of Atlas of Living Australia (Alluvium, 2016) and the Natural History Museum, London (Popov et al, 2021). Another area of pertinent importance is biodiversity discovery and related activities, for which a cost-benefit analysis was performed for Australia's species (Deloitte Access Economics, 2020). We will briefly review these studies, starting with the latter.

4.1. Cost benefit analysis of a mission to discover and document Australia's species

The Australian Academy of Science launched in 2021 a 25-year mission called Taxonomy Australia, with the goal to discover all remaining Australian species in a generation. To support this strategic plan, a cost-benefit analysis found that every AUD $1 invested in discovering all remaining Australian species would bring up to $35 of economic benefits (Deloitte Access Economics, 2020). The rapid analysis estimated that a total cost of 824 AUD over a period of 25 years would result in benefits of 3.7 to 28.9 billion AUD. To define scenarios, the study considered three levels for their calculations: a high change, a base and a low change.

For the analysis, the benefits of four major areas were estimated:

- **Biosecurity**: this sector considers threats by exotic invasive species that threaten Australia's biosecurity, native species and environment. It also considers non-genuine threats corresponding to suspected detections that are later confirmed to pose no or low risk. The benefit would result from the early detections and avoidance of misidentifications, which reduces delays for reaching taxonomic certainty and diagnosis. In the case of genuine threats, the rate of successful detection would result in a threat every 5 years (base scenario), one every 10 years (low change scenario) and one every 15 years (high change scenario). For non-genuine threats, the impact would result in avoiding them to one every 5 years (base scenario), one every 10 years (low change) and one every 15 years (high change);
- **Biodiscovery**: the benefits of more cost-effective and strategic testing of samples for drug discovery, in the research, pre-commercial phase, and of subsequent health benefits. There will be an increase of the biodiscovery value chain, resulting from agreements and contracts between researchers and pharmaceutical companies. At the stage of development, the increase in benefits results from the increase of

successful commercialization and sales, due to a larger pool of base species, and a more targeted species sampling. Finally, additional benefits result from avoided deaths attributed to prescription of natural product-based drugs and medicines;

- **Agricultural R&D**: the benefits of agricultural R&D would result at several levels, including increased knowledge of soil bacteria species that enhance crop management, soil fertility, and harmful organisms such as nematodes, or use of non-agricultural species in the transition to non-farm production of protein and carbohydrates; knowledge about crop wild relatives resulting in better resistance of crops to threats or trait benefits;
- **Biodiversity conservation**: the benefits of improved conservation outcomes with better informed decision-making, through a better understanding of species and their role within a given ecosystem. This includes promoting species resilience and strengthening ecosystems against environmental stressors. Furthermore, the goal of preventing extinctions is well understood by the public.

In addition to these areas, the report mentions other aspects of potential benefits of taxonomic discovery which were not considered including tourism, human and animal health, biomimicry, environmental monitoring and other sectors.

4.2. Atlas of Living Australia’s Impact and Value

The report of the assessment of the Atlas of Living Australia’s Impact and Value was performed in 2016 (Alluvium, 2016). The Atlas of Living Australia (ALA) is a RI supported by NCRIS, an Australian Government initiative, with the mission to provide free, online access to a vast repository of information about Australia’s biodiversity. The RI targets a major barrier resulting from the fragmentation and inaccessibility of biodiversity related data, generated and housed in museums, herbaria, collections, universities, research organisations, and government departments and agencies. ALA implemented a collaborative, digital and open infrastructure that aggregates biodiversity data from multiple sources, and focuses on making biodiversity information accessible and usable.

The evaluation exercise includes:

- an assessment of the key impact areas of the ALA such as influence on cultural change, new products and services, productivity and efficiency gains and applications and derivatives.
- initial and contemporary estimate of the benefit-cost ratio for investment in ALA and contextualising this in the organisation’s overall value.

The analysis considers information as an economic asset, which results in benefit by holding or using it. In the case of information as an economic asset, in relation to other assets, the following specifics apply (according to Moody and Walsh, 1999):

- information is infinitely shareable, reusable and repurposable;
- the value of information increases with use;
- information is perishable;

- the value of information increases with accuracy;
- the value of information increases when combined with other information;
- more is not necessarily better;
- information is not depletable.

The Theory of Change approach was used as methodology for the analysis, as depicted in Figure 2, extracted from the report (Alluvium , 2016).

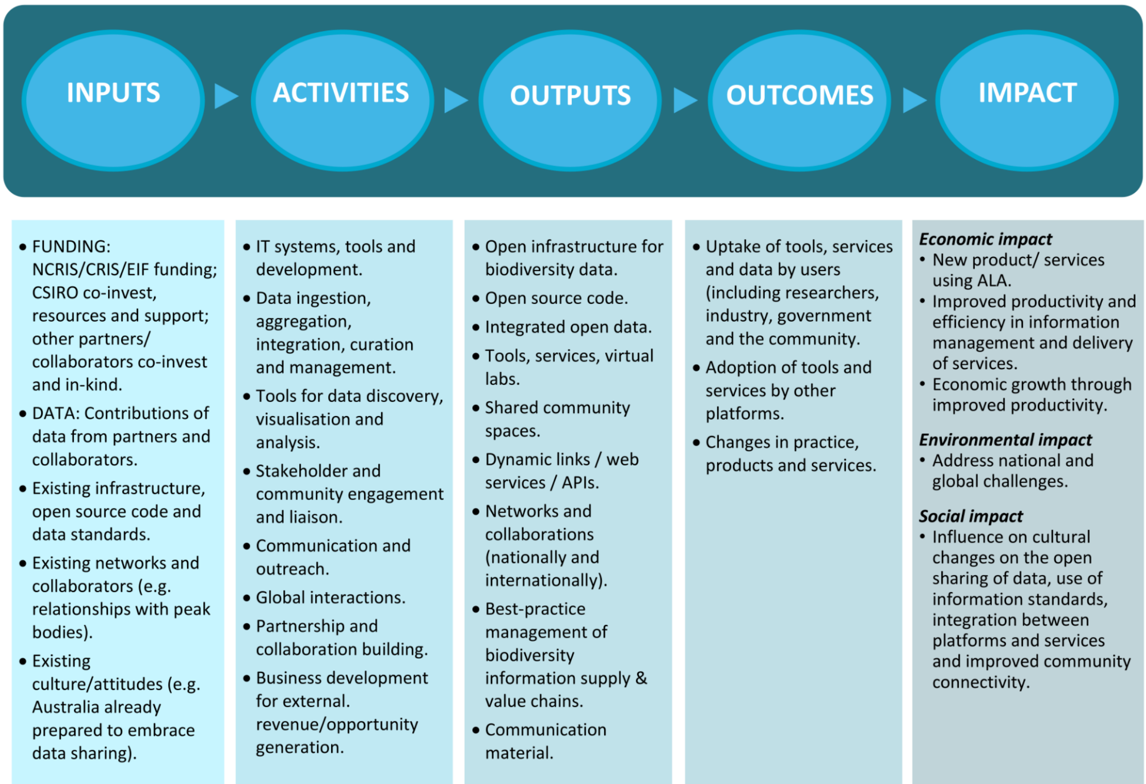

Figure 2. Impact pathway applied in the assessment of ALA RI (Alluvium , 2016).

The assessment exercise, based on online surveys, individual interviews, web metrics and case studies, was developed for two output areas and five impact areas (Table 6).

Table 6. Output and impact areas of the assessment of ALA.

Output area	Number of indicators	Type of indicators
Data	1	quantitative
Tools, services and infrastructure	1	quantitative/narrative
Impact area		
Influence on Cultural Change	6	quantitative/narrative

New Products and Services	3	quantitative/narrative
Productivity and Efficiency	5	quantitative/narrative
Applications and Derivatives	4	quantitative/narrative

ALA has led to a range of delivered and potential impacts, including: increased open sharing of data and standards; production of reports, papers and publications; significant efficiency gains for biodiversity data management and on-ground intervention and actions relating to biodiversity. The ALA Impact Evaluation indicated efficiency gains applied to Commonwealth expenditure on biodiversity and national parks to be 26.9 million AUD in 2016, with a benefit-cost ratio of 3.5:1.

4.3. The Value of Digitising Natural History Collections

A study commissioned by the Natural History Museum, London, aimed to determine the economic impacts of the digitisation of the 80 million specimens held in collections (Popov et al., 2021). In the scope of the study, digitisation may include several processes, like data transcription to databases, imaging, microscopy and computerised tomography scans, chemical, and molecular or genomic analyses.

Digitisation may result in several benefits, related to the increase of accessibility of collections, which become available: i) to a global audience at a lower cost, compared to in-person visits; ii) to the searchability of data transcribed or extracted, including its integration with other data; iii) to the preservation of specimens, for which physical handling requests will be lower and cause less damage, and iv) to the interaction of researchers with the collection, not limited by physical space or time, enabling multiple accesses to specimens.

The study applied a methodology based on a theory of change/logic model, which used inputs from museum collaborators and literature review to identify different pathways to impacts or benefits, how these will be materialised, their significance and who will benefit (e.g. visitors, scientists, taxpayers, society at large). The model developed is reproduced in Figure 3.

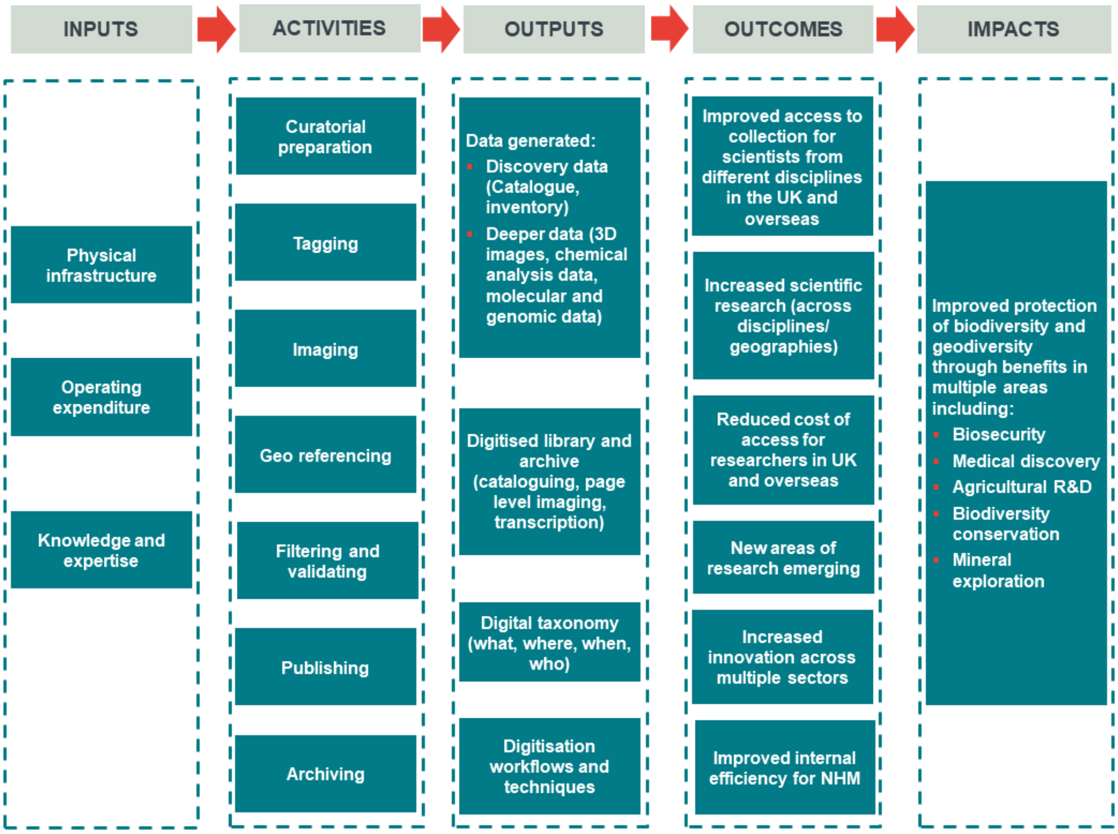

Figure 3. Theory of change showing the four components (inputs, activities, outputs and outcomes) with examples that lead to impact (Popov et al. (2021), https://doi.org/10.3897/rio.7.e78844.figure7).

The analysis took two approaches in valuing the impact of digitisation, namely on the return of investment:

- **top down** - estimation at the aggregate level of the expected returns an investment in science is likely to generate. This includes cost savings in terms of researchers not having to travel, or the amount of new research made possible;
- **thematic** - valuing specific benefits in a particular research area expected from digitisation, in five thematic areas - biodiversity conservation, invasive species, medicines discovery, agricultural research & development, and mineral exploitation (Figure 4).

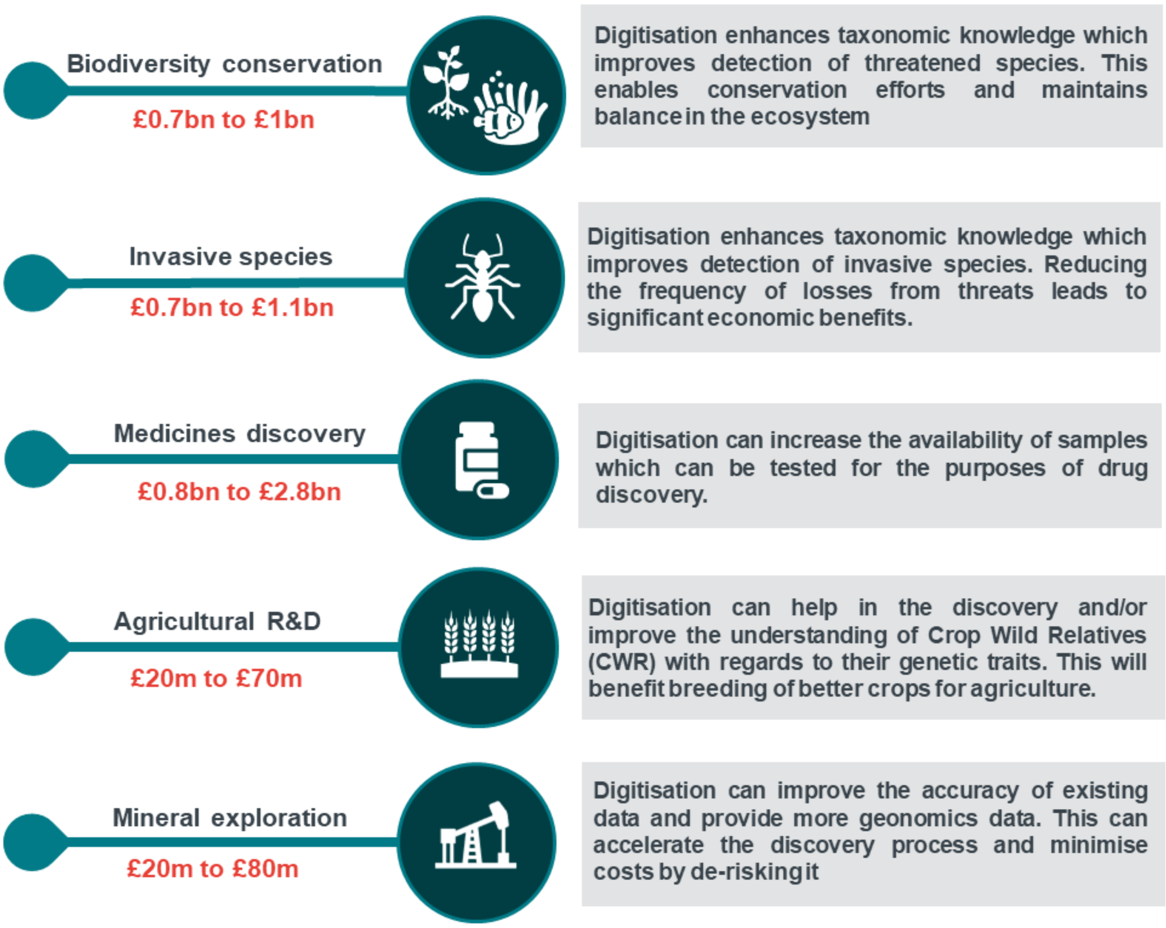

Figure 4. Valuing pathways to impact across five key areas (Popov et al. (2021), https://doi.org/10.3897/rio.7.e78844.figure3).

For the thematic approach, and while acknowledging limitations of data available, the study applied the estimates in Table 7.

Table 7. Economic benefits and estimates of the thematic approach to the valuing study of collections digitisation (Popov et al, 2021).

Thematic area	Economic benefits	Estimates
Biodiversity conservation	Efficiency of identification of threatened species Reduction of information gaps for countries rich in biodiversity but poor in biodiversity data	Estimate the value UK citizens place on preventing species declining anywhere in the world; Estimate the rate at which digitisation accelerates the identification of threatened species.
Invasive species	More comprehensive and updated database to identify	Estimate the reduction in time by avoiding delay/uncertainty in detecting

	species. Faster and easier access to specimen databases	threats; Estimate the reduction in damages due to the faster detection of threats or prevention of misdiagnoses, with greater certainty.
Medicine discovery	Research value - biodiscovery of collection samples with the potential to be explored for bioactive compounds Development value - commercialised value of the species sample	Estimate the impact of digitisation on the number of samples available for testing; Estimate the health benefits due to the increased number of commercialised samples.
Agricultural R&D	Accelerate the rate at which researchers are able to discover and improve their understanding of different natural species Faster and easier access to crucial information that can speed up the research process	Identify the rate at which digitisation increases the discovery and/or understanding of natural species for the purposes of agricultural R&D; Identify how this increased research creates economic value.
Mineral exploration	Efficiency of discovery Efficiency of processing	Estimate how digitisation affects the discovery process and/or fundamental scientific research; Estimate the value of any efficiencies it helps to achieve during discovery and/or try to value the fundamental research that might take place.

05 SCOPE OF DISSCO AND AREAS OF SOCIO-ECONOMIC IMPACT

The DiSSCo RI is a distributed research infrastructure in the domain of environment that aims to create a new business model to integrate all European natural science assets under common access, curation, policies and practices, for one European collection. It will:

- support and coordinate the digitisation of data, under the FAIR principles framework, across the distributed facilities;
- promote the necessary developments (scientific, technical, social) to link dispersed scientific information in NHC derived from the study of scientific collections;
- enable an unique access point for integrated data analysis and interpretation through a wide array of digital services provided by its community;

DiSSCo was included in the ESFRI Roadmap 2018, which Landscape Analysis identified needs related to the taxonomic gap and impacts of invasive alien species to biodiversity (ESFRI 2018). The services and activities of DiSSCo are rooted in the already multi-century-old representation of the earth's life and natural resources, through natural history collections based in museums, botanical gardens and universities.

The definition of a SEI analysis of DiSSCo must have in consideration the type of infrastructure, its phase in the life cycle of an RI and, with special focus, the scientific domain of the community it will provide services and its strategic objectives. Kolar et al (2019) analysed the relevance of the KPIs proposed by the ESFRI WG (ESFRI 2019) in relation to its scientific domain. They found that there are significant differences in the relevance of certain indicators depending on the ESFRI domain of the RI. The indicators need to be adapted to the type of infrastructure, and further work with the involvement of RI in the domains of energy, environment and health to ensure good monitoring.

DiSSCo is currently in the Preparation Phase, and according to its timeline,it is expected to complete implementation and become operational in 2026. The type and applicability of the indicators are naturally different for each phase, although some are applicable at all phases. This factor needs to be considered also when selecting indicators for the SEI set of DiSSCo.

Another dimension of impact strategically important for RI in the context of ESFRI projects and Landmarks is the participation in an “integrated ecosystem” that ensures links and complementarity between national and European priorities. The interconnected RI should promote frontier research, under an interdisciplinary paradigm. However, achieving this integration is not effortless, although it can also be a source of innovation. These aspects of integration - interoperability and connections with other RI in the same or related knowledge domains - are also to be considered in the monitoring schema developed.

5.1. Areas of impact of DiSSCo

The initial proposal of DiSSCo to ESFRI previews the impact of the RI, related to its goals, in several areas, which need to be captured by indicators. These include:

- **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 metabolomic)
  - new cultivars and animal breeds
  - new standards
  - new materials inspired by nature
- **Direct socio-economic impacts**
  - job creation
  - industry-oriented economic benefits
    - impact on organisations
    - applications in agriculture, environmental assessment, land use planning
    - new hardware/software - Small and Medium Enterprises
- **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
- **Innovation activity in the production of goods and services**
  - Direct contributions to food, textile, building materials, medicines, provision of sustainable energy, rare minerals, ecosystem services
- **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
- **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

5.2. Users of DiSSCo

The user groups of DiSSCo range from scientific researchers, to citizens and to decision makers. As identified by the user cases and user stories compilation (reports from Task 1.1 and 1.2, Fitzgerald et al., 2021, von Mering et al., 2021), the groups are:

- Research (academic, non-academic, including Citizen Science)
- Collection management
- Technical support (IT & IM)
- Policy (institutional, national & international)
- Education (academic & non-academic)
- Industry
- External (media & empowerment initiatives)

The services and support of these groups should be captured by the impact assessment. Many of these user groups are also stakeholders of the infrastructure. The interests of these vary, depending on their needs or different strategic visions. The short list of DiSSCo stakeholders and interests are presented in Table 8.

Table 8. Short list and interests of stakeholders in DiSSCo.

Stakeholder	Interests
RI funders (national, regional, others)	• Help to achieve the vision and mission in relation to the communities they represent/are interested in • Justify the investment
Scientific community	• Improved capacity to develop research activities
Industry	• Develop innovation products • Use of digital resources (images, videos) in paper (textbooks, newspapers and magazines) and digital products (webpages, platforms, etc.). • Science Tourism and Leisure
Other RI	• Leverage technological developments • Promote interoperability and integration at the scientific domain or regional level

RI management	• Monitor impact and achievements compared to goals, vision and mission
National nodes and institutional partners	• Improve and upgrade services to their users • Better coordination • Reach wider communities

5.3. Services of DiSSCo

New services to be or in implementation by DiSSCo and aligned projects will promote a transformation in the way users, particularly researchers, will access and use natural history collections. The digitisation of collections makes it possible to create services based on common access, curation, policies and practices. The set of planned services are identified in https://www.dissco.eu/services/, which is briefly listed here.

European Loans and Visits System (ELViS) – a web platform to provide a unified way to request visits, loans and virtual access. ELViS will enable digitisation on demand and support for collaborating on Virtual Access ideas and proposal submission. The request mechanism implemented in ELViS also enables future services for tracking usage metrics, monitoring and reporting and connecting collection usage with research outputs. This tracking service is important to support SEI indicators.

Collection Digitisation Dashboard (CDD) – a visual dashboard with information about the digitisation status, content and strengths of collections across the community of institutions. It displays progress in digitisation and provides summaries and comparisons regarding the number of objects, taxonomic scope, categories of preservation, stratigraphic age, geospatial range, level of digitisation and digital content availability for reuse. The data aggregation and compilation to support the dashboard can be used to support SEI indicators.

Specimen Data Refinery (SDR) – this will be another transformative service to be provided by DiSSCo, based on new models of digitisation workflows that process individual specimens and their metadata one-by-one into a model of industrial scale digitisation. It will integrate artificial intelligence and human-in-the-loop approaches to extract, enhance and annotate data at scale from digital specimen images and records (Walton et al. 2020). The refinery can enable potential applications by third party providers such as automated condition checking of specimens, natural language descriptions provision for specimens and taxonomic trait extraction.

Knowledge Base (KB) – a digital document repository of DiSSCo and related projects to support the infrastructure and users. It contains technical documentation and documented decisions, training materials, Frequently Asked Questions (FAQs), best practices, guidelines, and recommendations. The tools of the digital repository can be used to calculate indicators related to knowledge outputs.

Authorisation and Authentication Infrastructure (AAI) – this service will enable easy and integrated user authentication in the access to digital services, which can be linked to federated AAI services like eduGAIN, to enable institutional authentication or to external services like ORCID. This service will enable granularity of services authorization to access data that has legal restrictions, such as sensitive data on rare species. This service can be used also as a resource for SEI indicators, in the segmentation of DiSSCo users.

Unified Curation and Annotation System (UCAS) – the service will enable the curation and annotation on the Digital Extended Specimens (DS) for experts in the community and for machines. Transactions on the data will be stored as well as provenance information related to the curation or annotation events. The DS is a digital twin of the physical specimens that will link to data derived from the specimen (sequences, morphological data, taxonomic traits). SEI indicators can be derived from the number of transactions.

Digital Specimen Repository – this service will be implemented as a data repository for experimentation with Digital Specimen and other DiSSCo-related FAIR Digital Objects. It uses Cordra software to manage the digital objects and resolvable identifiers (Handles, DOI).

Self-assessment tool – this tool is intended to support teams, institutions and national nodes in developing organisational readiness for provision of the DiSSCo services and data, helping them to identify and target areas for improvement. The aim is for this to tie into the future provision of training and support, as well as helping to identify the gaps at aggregate level where that training may be most useful.

Helpdesk – a central place for all questions related to DiSSCo services or access programmes such as the virtual access and transnational access calls in ELViS. It will be integrated with DiSSCo services. The service will use JitBit software, a ticketing platform which can be a source of data for SEI indicators.

This collection of services will promote changes in the access and use of natural history collections, both internally in the institutions and staff linked to these resources, and externally to the researcher community that access collections and related services. The SEI indicators should be selected or defined in order to enable them to capture these changes and its impact.

06 COMPILATION OF SEI INDICATORS

Four sources were used to compile potential indicators to assess DiSSCo SEI. These include three frameworks specifically developed to assess the impact of RIs - ESFRI (2019), OECD (2019), Helman et al. (2020) - and an assessment study applied to the RI Atlas of Living Australia (Alluvium, 2016), which scope is closely related to DiSSCo. The file with the compiled indicators is included in Appendix 1, and is also available as a web data resource at https://tinyurl.com/DISSCO-SEIcompilation.

The review of these sources resulted in the construction of a table containing 37 descriptors (columns), and 210 transcribed indicators (Appendix 1, sheet "list_indicators"). Other assessments of SEI by RI were reviewed (Mirasgedis et al., 2018, 2019), but these did not provide significant new indicators as they used indicators already included in the previous reports, or specific indicators only applicable to the specific RI.

The full table contains many duplicated indicators, as it would be expected. Some duplications require careful analysis, because they may result from the breakdown of an indicator. Additionally, many indicators may be redundant by reporting similar impacts. The potential duplication and redundancy was signalled in the table column of related indicators, and also in the Appendix 1, sheet "related_indicators".

The completeness of indicator descriptors varies depending on the source. The ESFRI framework provided the most complete set of descriptors, which include definition, rationale, objective, detailed information about data needs, possible sources of data, indicator calculation, estimated costs for data collection, the level of the reporting burden, frequency of measurement and assumptions. For indicators from other sources, only part of these descriptors were available for each indicator. At this stage, no additional effort was made to add descriptors to those indicators, as these included the title, definition and rationale, which is sufficient to assess their relevance to DiSSCo. However, it will be necessary to revisit these descriptors for the final list of selected indicators, in order to complete them.

A consolidated table was prepared after the removal of duplications and columns which lack information for most of the sources (Appendix 1, sheet "list_consolidated"). This table retains 24 descriptors (columns) and 155 indicators.

Some descriptors, or in this case classifiers, were completed, regardless of the source framework, for all indicators that required some interpretation. The classifiers are:

- Type of indicator, based on RI-PATHS framework, with values: Activity, Outcome, Impact;
- Objective, based on ESFRI framework, see Table 2;
- Impact area, based on RI-PATHS framework, see Table 5;

- Impact category, based of OECD framework, see page 11;
- Nature of the indicator, with values: numeric, binary, categoric, narrative.

This exercise is useful for a general overview of the types of objectives and impact that the compiled list covers. Table 9 summarises the number of indicators that belong to each of the classifiers.

Additionally, the consolidated list contains 62 indicators of Activity, 54 of Impact and 39 of Outcome. The final set indicators to be used by DiSSCo should have a good balance of indicators in relation to the type of indicator, objective and category of impact, while considering the strategic objectives of the infrastructure, and the impact of services it will provide.

This **preliminary list of indicators** is a good basis for the identification of relevant indicators for the DiSSCo SEI assessment. However, gathering indicators from different sources revealed to be a challenge, because:

- there is no standard form of description of the indicators between frameworks;
- there is no or a lack of detail in the description of indicators by some frameworks, in relation to the rationale, possible sources of data gathering, indicator calculation, etc.;
- the definition of concepts might vary between frameworks;

Furthermore, this compilation exercise reveals that coverage of activities, outcomes and impacts to the economy or society, from indirect benefits of the environment, biodiversity conservation, or food security, to name a few, are sparsely covered by indicators. **Another example of an area lacking coverage is the impact of digitisation and digital access to services.** When covered, these topics are based on surveys to stakeholders, which is a costly method for data gathering, and which results may not be directly convertible into an indicator form, especially when **narrative responses** are gathered.

Table 9. Number of indicators of the consolidated table that fall into one of the classifiers of indicator type: objective, impact area and category of SEI impact. HR - Human Resources, E&I - Economy and Innovation.

		Objective									
		Delivery of education and training	Enabling Scientific Excellence	Enhancing Collaboration in Europe	Enhancing transnational collaboration in Europe	Facilitating economic activities	Facilitating international cooperation	Optimising data use	Optimising management	Outreach to the public	Provision of scientific advice
Impact area	Category of SEimpact										
HR	economic	1	1	0	0	7	0	0	0	0	0
	scientific	0	17	1	2	1	2	0	0	0	0
	technological	0	2	0	0	1	0	0	0	0	0
	training and education	15	0	0	0	1	1	0	0	0	0
E&I	economic	0	0	0	0	21	0	1	1	0	0
	scientific	0	1	0	0	0	0	4	0	0	0
	technological	0	2	0	0	14	0	9	4	0	0
Policy	scientific	0	0	0	0	0	0	0	1	0	0
	social and societal	0	0	0	0	0	0	0	0	1	16
	technological	0	0	0	0	1	0	0	0	0	1
Society	economic	0	0	0	0	0	0	0	0	1	0
	social and societal	0	0	0	0	0	0	0	5	15	2
	technological	0	0	0	0	1	0	2	0	0	0

07 NEXT STEPS

This report established the background for the preparation of a list of SEI indicators to be used by DiSSCo. That achievement will be reached through the following steps:

- Review the consolidated table of indicators to identify possible incoherences, lack of support information (e.g., definition, methods for calculation) and gaps;
- Assess applicability and preliminary relevance for DiSSCo, namely in terms of:
    - Indicators scope
    - Operationalization requirements
- Revise indicators (definition) to include specificities for DiSSCo
- Prepare a preliminary list of indicators for DiSSCo
- Perform a survey to DPP partners (WP leaders), including national nodes to assess the relevance of the indicators;
- Create a suggested table of SEI to be adopted by DiSSCo;
- Identify requirements of information and data sources for indicators;
- Provide Guidelines for the SEI of DiSSCo
- Provide guidance for future updates of indicators, namely to accommodate with recommendations on alignments with EOSCs KPIs (European Commission, Directorate-General for Research and Innovation, 2022).

08 GLOSSARY

Activity – Initiatives and endeavours undertaken using the resources of a Research Infrastructure or work performed by Research Infrastructure staff.

Activity indicator - Indicators that capture the scale and nature of a Research Infrastructure's activities; a measure that should form part of internal reporting.

FAIRification - informal term to designate the process in which data is transformed and framed by the technologies that enables them to be in accordance to FAIR (Findable, Accessible, Interoperable and Reusable) principles

Impact – Intended and unintended long-term effects of activities using the resources of a Research Infrastructure or work performed by Research Infrastructure staff.

Impact indicator - An indicator that reflects the extent and nature of generated effects in the economy and wider society; with few exceptions, impact indicators are estimations.

Outcome – Longer-term effects that stem from the stakeholder uptake of or interaction with Research Infrastructure outputs.

Outcome indicator - Indicators that document the result of the first productive interactions; collecting data by reaching out to involved stakeholders, e.g. via a survey, interview, external reporting or other data-gathering means.

09 ABBREVIATIONS AND ACRONYMS

ACTRIS - Aerosols, Clouds and Trace gases Research Infrastructure

ALA - Atlas of Living Australia

BCA - Benefit-Cost Analysis

CBA - Cost-Benefit Analysis

DiSSCo - Distributed System of Scientific Collections

EIA - Economic Impact Analysis

ERIC - Educational Resources Information Center

ESFRI - European Strategy Forum on Research Infrastructure

FAIR - Findable, Accessible, Interoperable and Reusable

KPI - Key Performance Indicator

MBPF - Marginal Benefit of Public Funds

MCA - Multicriteria Analysis

MCF - Marginal Costs of Public Funds

NCRIS - National Collaborative Research Infrastructure Strategy)

OECD - Organisation for Economic Co-operation and Development

RACER - Relevant, Accepted, Credible, Easy and Robust

RI - Research Infrastructure

RI-PATHS - acronym of project “Research Infrastructure imPact Assessment paTHwayS”

SEI - Socio Economic Impact

WTA - Willingness to Accept compensation

WTP - Willingness to Pay

10 CITED REFERENCES

Alluvium (2016). Assessment of the Atlas of Living Australia's impact and value. Report produced for the CSIRO. Canberra. 50 pp.

Brooke, M.L. (2000) Why museums matter. Trends Ecol Evol 15, 136–137

Deloitte Access Economics (2020). Cost benefit analysis of a mission to discover and document Australia's species | Deloitte Australia | Economics. Deloitte Access Economics Pty Ltd. 107 pp. https://www2.deloitte.com/au/en/pages/economics/articles/cost-benefit-analysis-mission-discover-document-australia-species.html

ESFRI (2018). Roadmap 2018. Strategic Report on Research Infrastructures. Part 2. Landscape Analysis, Environment. Printed on behalf of ESFRI by Dipartimento di Fisica – Università degli Studi di Milan. Available online at http://roadmap2018.esfri.eu.

ESFRI (2019). Report of the ESFRI Working Group on monitoring RIs performance. www.esfri.eu. (2019, December). https://www.esfri.eu/latest-esfri-news/report-esfri-working-group-monitoring-ris-performance

ESFRI (2021). ESFRI Strategic Report on Research Infrastructures Roadmap 2021 Public Guide. https://www.esfri.eu/sites/default/files/ESFRI_Roadmap2021_Public_Guide.pdf

European Commission, Directorate-General for Research and Innovation. (2022). Strategic Research and Innovation Agenda (SRIA) of the European Open Science Cloud (EOSC). https://doi.org/doi/10.2777/935288

Fitzgerald, H., Juslén, A., von Mering, S., Petersen, M., Raes, N., Islam, S., Berger, F, von Bonsdorff, T., Figueira, R., Haston, E., Häffner, E., Livermore, L., Runnel, V., De Smedt, S., Vincent, S., Weiland, C. (2021). DiSSCo Prepare Deliverable D1.1 Report on Life sciences use cases and user stories. https://doi.org/10.34960/xhxw-cb

Greco, S., Ehrgott, M., & Figueira, J. R. (Eds.) (2016). Multiple criteria decision analysis: state of the art surveys. (2nd ed.) (International Series in Operations Research and Management Science; Vol. 233). Springer. https://doi.org/10.1007/978-1-4939-3094-4

Hajdinjak, M. (2019). A Practical Guide: Assessment of Socio-Economic Impacts of Research Infrastructures—Project ResInfra@DR (A. Havas, Ed.). Centre for Social Innovation. http://www.interreg-danube.eu/uploads/media/approved_project_output/0001/32/0b698b634cc4cc3df75dd303bdc3917f5ed6e2af.pdf

Helman, A. Barberis, M., Vignetti, S., Catalano, J., Griniece, E., Kroll, H., Zenker, A., Martin C., (2020). Validated IA Model, deliverable 5.1 produced in the framework of H2020 project Research Infrastructure imPact Assessment paTHwayS (Grant agreement 777563). https://ri-paths.eu/wp-content/uploads/2018/03/D5.1_Validated-IA-model.pdf.

Johansson, P. O., & Kriström, B. (2016). Cost-benefit analysis for project appraisal. Cambridge University Press.

Kolar, J., Cugmas, M., & Ferligoj, A. (2019). Towards Key Performance Indicators of Research Infrastructures. ArXiv:1910.00304 [Cs, Stat]. http://arxiv.org/abs/1910.00304

Mirasgedis, S., Sarafidis, Y., Mihalopoulos, N., Georgopoulou, E., Mona, L., Papagiannopoulos, N. (2019). Deliverable 8.3: Report on case studies for the socio-economic impact of ACTRIS

Mirasgedis, S., Mihalopoulos, N., Sarafidis, Y., Georgopoulou, E., & Theodosi, C. (2018) Deliverable D8. 1: Report on KPIs for the quantification of ACTRIS direct impact.Moody, D. L., & Walsh, P. (1999). Measuring the Value Of Information-An Asset Valuation Approach. ECIS, 496–512.

OECD (2015), Data-Driven Innovation: Big Data for Growth and Well-Being, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264229358-en

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

Popov, D., Roychoudhury, P., Hardy, H., Livermore, L., & Norris, K. (2021). The Value of Digitising Natural History Collections. Research Ideas and Outcomes, 7, e78844. https://doi.org/10.3897/rio.7.e78844

Rocha L.A., Aleixo A., Allen G., Almeda F. et al (2014) Specimen collection: an essential tool. Science 344, 814–815

Ruttkay, Z., & Bényei, J. (2018). Renewal of the Museum in the Digital Epoch. In G. Bast, E. G. Carayannis, & D. F. J. Campbell (Eds.), The Future of Museums (pp. 101–116). Springer International Publishing. https://doi.org/10.1007/978-3-319-93955-1_10

Suarez A.V., Tsutsui N.D. (2004) The value of museum collections for research and society. Bioscience 54, 66–74

Tewksbury J.J., Anderson J.G.T., Bakker J.D. et al (2014) Natural history's place in science and society. Bioscience 64, 300–310

von Mering, S., Petersen, M., Fitzgerald, H., Juslén, A., Raes, N., Islam, S., Berger, F, von Bonsdorff, T., Figueira, R., Haston, E., Häffner, E., Livermore, L., Runnel, V., De Smedt, S., Weiland, Claus (2021) D1.2 Report on Earth sciences use cases and user stories. https://doi.org/10.34960/n3dk-ds60

Vignetti, S., Griniece, E., Cvijanovic, V., Reid, A., Helman, A., & Kroll, H. (2019). Deliverable 3.2 Title: Report on stocktaking results and initial IA model. RI-PATHS Consortium.

Vignetti, S. (2021). Designing a Research Infrastructure with Impact in Mind. In H. P. Beck & P. Charitos (Eds.), The Economics of Big Science: Essays by Leading Scientists and Policymakers (pp. 79–84). Springer International Publishing. https://doi.org/10.1007/978-3-030-52391-6_11

Walton, S., Livermore, L., Bánki, O., Cubey, R.W.N., Drinkwater, R., Englund, M., Goble, C., Groom, Q., Kermorvant, C., Rey, I., Santos, C.M., Scott, B., Williams, A.R., Wu, Z. (2020). Landscape Analysis for the Specimen Data Refinery. Research Ideas and Outcomes 6: e57602. https://doi.org/10.3897/rio.6.e57602

11 OTHER REFERENCES

Barker, M., Olabarriaga, S. D., Wilkins-Diehr, N., Gesing, S., Katz, D. S., Shahand, S., ... & Costa, A. (2019). The global impact of science gateways, virtual research environments and virtual laboratories. Future Generation Computer Systems, 95, 240-248. https://doi.org/10.1016/j.future.2018.12.026

Beagrie, N., & Houghton, J. (2016). The value and impact of the european bioinformatics institute.https://vuir.vu.edu.au/33707/1/EBI-impact-report.pdf

Clough, G. W. (2013). Best of both worlds: museums, libraries, and archives in the digital age. Smithsonian Institution. https://repository.si.edu/handle/10088/21153

Ellenberg, J., Swedlow, J. R., Barlow, M., Cook, C. E., Sarkans, U., Patwardhan, A., ... & Birney, E. (2018). A call for public archives for biological image data. Nature Methods, 15(11), 849-854.https://doi.org/10.1038/s41592-018-0195-8

European Commission Directorate-General for Regional and Urban policy (2014). Guide to Cost-Benefit Analysis of Investment Projects - Economic appraisal tool for Cohesion Policy 2014-2020. https://ec.europa.eu/regional_policy/sources/docgener/studies/pdf/cba_guide.pdf

Fell, M. J. (2019). The economic impacts of open science: A rapid evidence assessment. Publications, 7(3), 46.https://doi.org/10.3390/publications7030046

Florio, M. (2019). Investing in science: Social cost-benefit analysis of research infrastructures. MIT Press.

Gabella, C., Durinx, C., & Appel, R. (2017). Funding knowledge bases: towards a sustainable funding model for the UniProt use case. F1000Research, 6.

Martin, C. S., Repo, S., Arenas Márquez, J., Blomberg, N., Lauer, K. B., Pérez Sitjà, X., ... & Smith, A. (2021). Demonstrating public value to funders and other stakeholders—the journey of ELIXIR, a virtual and distributed research infrastructure for life science data. Annals of Public and Cooperative Economics, 92(3), 497-510.

Miller, C. G., Brewer, P., Carine, M., Comerford, G., Hardy, H., Hart, A., ... & Woodburn, M. (2022). Join the Dots: assessing a collection of 80 million items at The Natural History Museum, London. Museum Management and Curatorship, 1-20.

Nelson, G., & Ellis, S. (2019). The history and impact of digitization and digital data mobilization on biodiversity research. Philosophical Transactions of the Royal Society B, 374(1763), 20170391.https://doi.org/10.1098/rstb.2017.0391

Reid, A. (2021). Designing a socio-economic impact framework for research infrastructures: Preliminary lessons from the RI-PATHS project. In The Economics of Big Science (pp. 61-70). Springer, Cham. https://library.oapen.org/bitstream/handle/20.500.12657/42933/2021_Book_TheEconomicsOfBigScience.pdf?sequence=1#page=65

I. Appendices

Appendix 1. Table of compiled indicators of socio-economic impact, available as a web data resource at https://tinyurl.com/DISSCO-SEIcompilation.