ENLIGHT RISE- RESEARCH AND INNOVATION AGENDA WITH AND FOR SOCIETY: LEVERAGING DIGITAL INNOVATION FOR A GREENER AND HEALTHIER EUROPE

WP No	Del. Rel. No	Del No	Title	Lead beneficiary
3	D3.1	D17	Mapping of Digital Research Infrastructural Resources	UU

Nature	Dissemination Level	Related to Del. No (if applicable)
Report	Public	

Description (short)
The report brings together a mapping of Digital Research Infrastructural Resources across ENLIGHT RISE partners. The WP 3 mapping exercise aims at creating an overview of platforms, competencies and priorities within the ENLIGHT community. The objective is to identify synergies related to the ENLIGHT flagship areas and provide input for future joint R/I actions.

Document Version Control		
Version 0.3 DRAFT AWAITING EUROPEAN COMMISSION APPROVAL	Originated by: Margaretha Andersson, Liisi Lembinen, Birgit Schmidt, Merle Schatz	UU, UT, UGOE

Table of contents

Executive Summary

Access to world-class digital research infrastructure is a cornerstone of a competitive ENLIGHT RISE R&I agenda. The goal of this report is to map existing and identify needs of access to digital research infrastructures across the alliance. Initially, a mapping exercise was performed to gather information about the present situation within the alliance. The situation turned out to be promising, with many existing digital platforms in wide range of scientific areas and a strong engagement in a majority of the partner universities. Since innovation is a scope of the ENLIGHT community there was also a smaller questionnaire targeting researchers in Uppsala using "sensitive data". Researchers expressed needs especially in areas as computing capacity, storage and legal advice. However, a drawback was that many researchers have only a faint idea about the size of their needs. Readiness to take part in governance was also high with several representatives form ENLIGHT partners already active in steering committees, EOSC task forces and national governing bodies.

On basis of the mapping questionnaire, satisfying coverage was demonstrated in relation to the European landscape of digital research infrastructures. ENLIGHT connections could be identified within all five ESFRI Science Clusters. There was also substantial engagement in EOSC and resources related to computing capacity and high performance computing (HPC, referring to large scale resources for computing (supercomputers) and storage able to solve complex problems requiring massive computation capacity). Projects in a wide range of scientific disciplines already demonstrating best practise related to data driven science were also identified. The ENLIGHT community was then concluded to be equipped enough to work with connecting and sharing within existing major European processes and models for connecting and sharing. On the basis of the mapping exercise, a Roadmap (visual strategic plan) to ensure access to digital infrastructural resources within the ENLIGHT Alliance was finally presented.

Abbreviations

ENLIGHT partner	Short Name	Country
University of Bordeaux	UBx	FR
Ghent University	UGent	BE
Comenius University Bratislava	CU	SK
University of the Basque Country	UPV/EHU	ES
National University of Ireland, Galway	NUIG	IE
University of Göttingen	UGOE	DE
University of Groningen	RUG	NL
University of Tartu	UT	EE
Uppsala University	UU	SE

AI – Artificial Intelligence
DI – Digital Infrastructure
DRI – Digital Research Infrastructure
EEA – European Education Area
ERA – European Research Area
ENVRI – ESFRI/EOSC Cluster for Environmental Research Infrastructures
ESCAPE – ESFRI/EOSC Cluster of Astronomy and Particle Physics
ESFRI – European Strategy Forum on Research Infrastructure
EOSC – European Open Science Cloud
HE – Higher Education
HEI – Higher Education Institution
HPC – High Performance Computing
LIFE – ESFRI/EOSC Cluster for research infrastructures in life science
PaNOSC – ESFRI/EOSC Photon and neutron Open Science Cloud
RI – Research Infrastructure
SSHOC – ESFRI/EOSC Social Sciences and Humanities Open Cloud

INTRODUCTION

ENLIGHT is a European University formed by nine comprehensive, research-intensive universities from nine European countries (Belgium, Estonia, France, Germany, Ireland, Netherlands, Slovakia, Spain, Sweden), collectively training over 300,000 learners per year and sharing a deep commitment to their social responsibility.

ENLIGHT aims to undertake a fundamental transformation of European higher education that empowers learners as globally engaged citizens with state-of-the-art knowledge, skills, and innovation potential to tackle major societal transitions and to promote equitable quality of life and sustainability. ENLIGHT wants to establish the foundations of an open, integrated European University System to ensure the free movement of students and staff including sharing of resources.

Since "connecting" and "sharing" are core values for ERA, EEA and the ENLIGHT-platform, access to digital resources for education, research and innovation are of crucial importance. The long-term objective is to create an ENLIGHT digital DI/AI ecosystem where academia and society may collaborate via digital RI resources.

The context of ENLIGHT and ENLIGHT-RISE

ENLIGHT (https://enlight-eu.org/) is a European university Network to promote equitable quality of Life, sustainability and global engagement through Higher Education Transformation. It brings together nine comprehensive, research-intensive, flagship universities with a strong reputation: Ghent University (UGent), University of Bordeaux (UBx), Comenius University Bratislava (CU), National University of Ireland Galway (NUI Galway), University of Göttingen (UGOE), University of the Basque Country (UPV/EHU), University of Groningen (RUG), University of Tartu (UT), and Uppsala University (UU). All share a deep commitment to their social responsibility.

ENLIGHT strives to transform the way in which we address global societal challenges by developing new models and methodologies for education and research adapted to the complex sustainability issues that cities and communities face today, focusing on five flagship challenges:

- Health and Well-being
- Digital revolution and the Impact of Digitalization
- Climate action
- Energy transition and Circular economy
- Equity

The ENLIGHT "Research and Innovation agenda with and for Society" (RISE) is a three-year effort, establishing foundations and initiating implementation of a common research and innovation (R&I) agenda within the ENLIGHT European University Alliance. The objective of the present ENLIGHT-RISE work package associated with this report (WP 3) is to enable the sharing of and optimize capacity of digital research infrastructures, while benchmarking and optimizing the societal impact of digital innovation/artificial intelligence (DI/AI) on the flagship challenges.

Access to world-class digital research infrastructures are pivotal to making ENLIGHT attractive for the best researchers, and crucial for deploying an ambitious joint R&I agenda. ENLIGHT-RISE will thus map digital research infrastructures in partner universities and propose a roadmap for sharing these and for connecting with European digital infrastructures for R&I. This will complement the ENLIGHT Interconnected Campus established under Erasmus+, which has an initial focus on making IT systems interoperable and connecting existing infrastructures for learning and administration related to education, making resources and services accessible across ENLIGHT.

Definition of digital research infrastructures

According to the EU "Research infrastructures are facilities that provide resources and services for the research communities to conduct research and foster innovation in their fields".[1] They can be used beyond research e.g. for education or public services and they may be single-sited, distributed, or virtual.

They include:
- major scientific equipment or sets of instruments,
- collections, archives or scientific data,
- computing systems and communication networks,
- any other research and innovation infrastructure of a unique nature which is open to external users.

"Open to external users" also include services, support personnel and often also processes to foster best practise and promote scientific excellence among users. Funding models depend on size and disciplinary research culture. In general, local infrastructures may be accessible for a fee – larger resources are more often governed by priorities related to scientific excellence.

In this report Digital Research Infrastructure/DRI are defined as sustainable digital platforms for mainly academic use – they may be computational resources, storage and digital tools/algorithms/software made openly available for users. Support for users is also available. Experimental equipment and non-digital resources are not a scope of this mapping (nor are digital platforms for administrative use). However, experimental equipment and e.g. biobanks

---

[1] https://research-and-innovation.ec.europa.eu/strategy/strategy-2020-2024/our-digital-future/european-research-infrastructures_en

and on-site resources produce data which enter the system of DRI when analytics etc. is performed. They are then a valuable source of information for further sharing, processing and discovery in projects related to research and innovation.

Digital Research Infrastructure (DRI) in the Life Cycle of Research Data

The concept of DRI is closely related to the research data life cycle as described in the figure below, where resources associated with an active research/innovation project are found on the right side on the circle and open science practices and resources are found on the left side. In the ENLIGHT context work package/WP 3 is focussing on DRI and resources as capacity for computing/storage, software and support for analytics while WP 7 focus on Open Science, Data Management and making research data FAIR (Findable, Accessible, Interoperable and Reusable). However, these two aspects must be coordinated in order to make e-science and digital innovation happen within the ENLIGHT platform as well as with society and stakeholders.
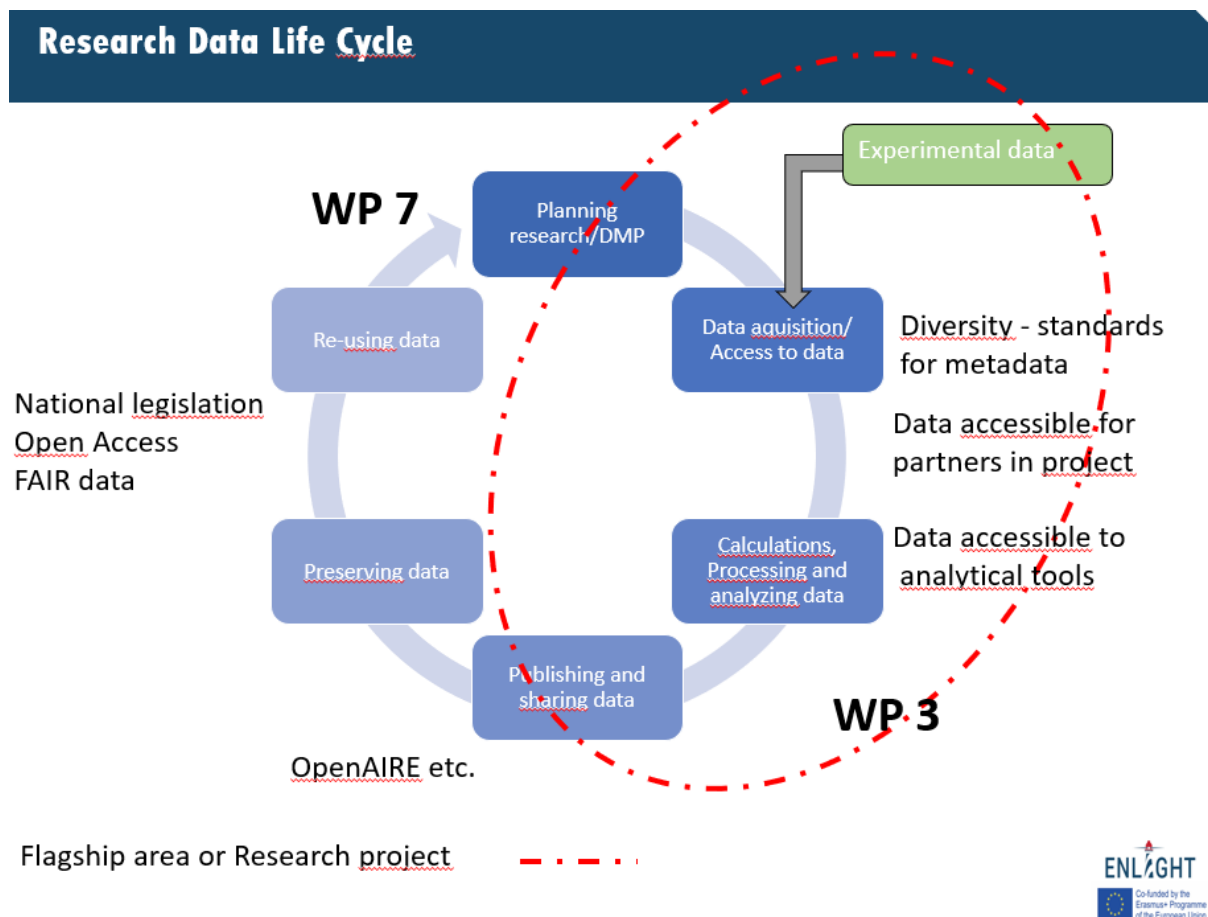
Support for Open Science in general is in all ENLIGHT partners associated with university libraries while DRI:s can be found anywhere in the organizations – often close to the researchers/departments or organized as infrastructural core facilities/platforms. These

differences also reflect the fact that research infrastructure in general are more of bottom up-driven resources operating close to research projects while Open Science relies stronger on aspects of harmonizing policies and practises, thus forming a framework related to information management and scholarly communication before/after a research project. An introduction to FAIR principles is available via OpenAIRE.[2]

Funding and prioritization

Even though a RI ideally is intended to be open for external users, constraints on external access may be a reality. Data are often shared for free when they are already published and accessible via an open service – however, when comes to computing resources/storage, software or support to get useful information out of huge data sets there are often some kind of "gate keeping". At this stage funding models and principles for priority when resources are limited also have to be addressed – it is a complex area regulated by a number of mechanisms:

- Many users are interested in RI:s with resource limitation
- Some resources are open for a fee – others may be for free. The latter are often free to some extent and for general use of limited size. If needs are more demanding – payment is required.
- Some skills and preparations are compulsory to be accepted as user – this is usually obtained by a local platform where training is organized.
- Large resources are invested to set up a RI – organizations may have made an investment on the basis of their needs. They expect priority when resources are limited. Ownership may be divided between academia and national bodies with different priorities.
- Competition is sometimes very hard – access may be granted on basis of scientific excellence but free for the prioritized groups/projects. Processes for application and evaluation may be time-consuming and no quick-fix even for those with excellent projects. Investments may be made at earlier stages by the research group in order to prepare for access (pilot projects, travels etc.).

All in all, the first part of this report will present information about DRI resources and support available within the ENLIGHT Alliance, but in order to actually use it there may be a number of boxes to tick:
- funding,
- training/experience from similar techniques,
- project of scientific or societal interest and
- interest from intended partners etc.

This is in itself a process of co-creation where ENLIGHT may provide an arena where stakeholders can meet.

---

[2] https://www.openaire.eu/how-to-make-your-data-fair

MAPPING EXERCISE and DATA COLLECTION

To create an overview illustrating the present status of digital research infrastructures (DRI) that are currently available within the ENLIGHT Alliance, two surveys were performed.

1) A mapping exercise completed by all nine ENLIGHT partners.
2) A smaller questionnaire targeting individual researchers in Uppsala focussing on "sensitive data".

Objective and Mapping Exercise

A questionnaire (implemented in Google Forms) mapping engagement and presence of resources in important DRI key areas was distributed to all ENLIGHT partners. The intention was to reach out to a limited group of specialists/research support officers to obtain overarching information about the DRI maturity on the institutional level.
The questionnaire was sent to the institutional contacts for the present work package (WP 3) with an option to involve more people – researchers, managers or supporting officers - when needed.

Since one main goal of the ENLIGHT Alliance is to foster harmonization and collaboration between universities, societies and countries it is inherent that collaboration may start from very different positions. It turned out to be the case in this mapping exercise. While some of the ENLIGHT partners are already very active on the international digital arena others have just started their journey. Initially, some contact persons were engaged in more than one WP and neither very familiar with the concept of DRI, nor with the interface between local and wider DRI-landscape. In some cases, the "close to researcher"-nature of DRI also appeared late in the process and partners discovered that there was a need to reach out to researchers to find out more about strengths, priorities, strains, demands and limitations in their own local community before the mapping exercise could be completed. The mapping process has thus been an *iterative* six-month-process with more than one delivery of information from several partners, personal contacts and dialogue about definitions etc.

By mid-June 2022 the questionnaire has been completed by the nine ENLIGHT-partners. Input was delivered by a centrally located officer responsible for coordinating partners-activities in digital research infrastructure related to ENLIGHT-RISE WP 3. In order to complete the form institutional contacts has been handled through WP 3 institutional contacts and may involve more people according to the organization of each HEI. Primary data will be handled according to the ENLIGHT Data Management Plan, including information about respondents.

In addition, a smaller survey targeting researchers in Uppsala using "sensitive data" with questions about collaboration partners and need for resources was performed. These answers are anonymous and the full report archived in Uppsala university.

RESULTS - Mapping of Digital Research Infrastructural Resources

The questionnaire was divided into seven target areas of major importance for the landscape of Digital Research Infrastructure. For each area it was possible to add answers in free text in order to target resources related to ENLIGHT flagship areas. The results are presented below.

1) Approaches for finding and getting access to research data at ENLIGHT universities

Is there a Data Office/Data Access Service at the University?

9 svar
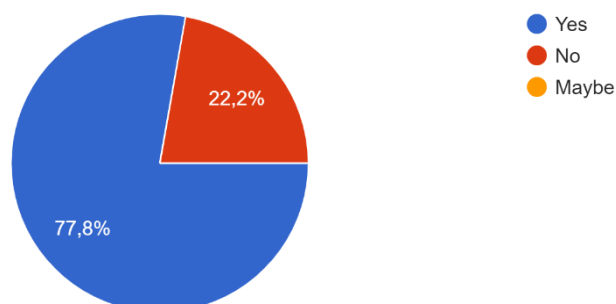
Answers from all nine partners indicate that there in most cases are some kind of coordinated "Data Office" supporting data management and access to already published data in a university. Relationships between libraries and many digital research infrastructures are strong. However, there are major differences in the maturity of the "data office". In some cases these are well established organizational functions, others have a network of dedicated individuals and some have more of a set of guidelines/documents to support users but not dedicated personnel. ENLIGHT-RISE WP 7 (Open Science) has made a survey to map "Open Science Readiness" within ENLIGHT. .

A research data contact point/data office is a long-term goal expressed individually by ENLIGHT universities. There is no established definition of a *data office* but there is an awareness of future needs and the ENLIGHT partners are all demonstrating a positive will to take on the journey. Some kind of coordination is available in all partner HEIs. A data office is however no guarantee that data is compliant to the FAIR principles. At this stage it must be stated that "how" is a concept of great diversity. Ambitions are for the future and does not include a systematic method to find and share data "in progress", nor already published "un-FAIR" data. There is also a challenge with differences in the quality of data to be aware of.

2) Digital research infrastructures (DRI) of critical mass and sustainability

"Critical mass and sustainability" are defined in the questionnaire as a platform with organizational support and funding including personnel ready to assist external users. It may be single sited or distributed. RI are often organized and funded in levels from local/ institutional resources via national platforms to larger international/globally available research environments. This organizational structure is applied in the questionnaire with answers as below.

The size of the platform is normally reflecting the critical mass and organizational maturity. National and international resources often provide some kind of broader responsibility for a research community of users thus operating upskilling activities etc. To enter the larger facilities there has to be a local contact/resource where newcomers in an area can embark.

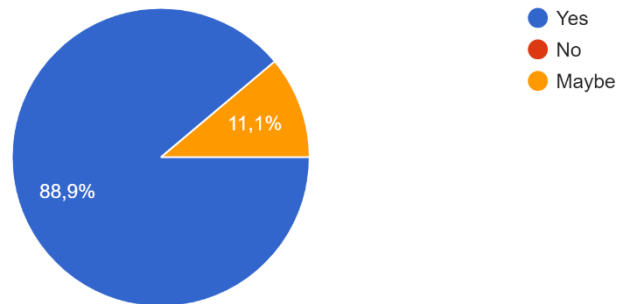On local level?

9 svar

- Yes
- No
- Maybe

11,1%

88,9%
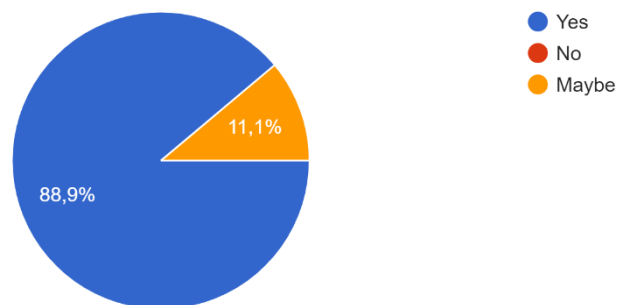
Partners offering high performance computing infrastructure: UGENT, UT, UC, UU, UBx
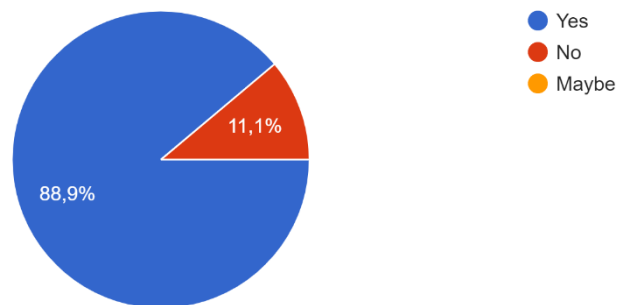Subject specific data centers are also available in some HEI: UGENT, NUIG, UU

Is the University actively involved in a European/ESFRI DRI as host or otherwise?
9 svar

- ● Yes
- ● No
- ● Maybe

88,9%   11,1%

Are You host or partner in a national DRI platform/consortium?
9 svar


- ● Yes
- ● No
- ● Maybe

88,9%   11,1%

As an overview, the answers indicate a high degree of involvement in DRI among the ENLIGHT partners. However, a "yes" may reflect a great diversity in local conditions including size/resources and disciplinary focus. There is no established definition used – answers reflect the informants' view. There may be one flagship area in a narrow discipline or a broad range of resources (hardware, software and personnel) in many disciplinary areas – few users or many. It is however an observation that there is also an awareness of the structural nature of infrastructure with involvement on more than one organizational level. To be accepted as host or partner in a national/international platform there must be local resources of sustainability, size and quality – already *recognized by the scientific community of users*. Within this survey no assessment of size/quality was made since it is already made by the scientific community of users. It also indicates a readiness to take part in governance – at least to some extent.

3) Relationships between institutional strategies and involvement in DRI, or DI/AI

Is there a relationship between institutional strategies and decision to involve in DRI and/or DI/AI?
9 svar

- Yes
- No
- Maybe

11,1%
11,1%
77,8%
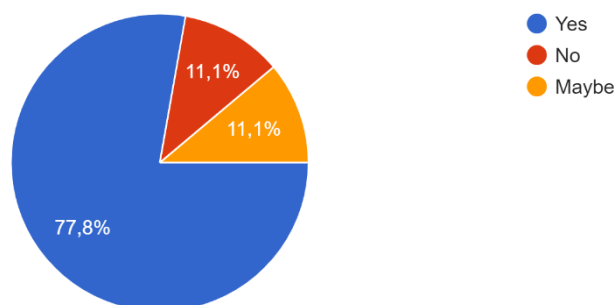
In general, institutional strategies are linked to local areas of research excellence. Answers illustrates an institutional awareness about the importance of digital resources in HE. They also fit in well with the ENLIGHT Flagship Areas. Sometimes DRI is described as an area of its own in the strategic plan of a HEI– in other cases, access to any kind of research infrastructure is highlighted as an area of great strategic importance. In these cases, digital components are considered to be a part of a larger infrastructural landscape with non-digital *and* digital resources. Of course, in many cases they cannot be separated. There is no research without research data – and experimental work often produces large amounts of experimental data to share, analyze and store.

4) Projects demonstrating "best practice" in data driven research and innovation;

A range of examples demonstrating "best practice" associated with the uses of digital infrastructural resources or services/expertise were highlighted by ENLIGHT partners. Results are presented in the Appendix, as a table where strengths in DRI and areas/projects demonstrating best practice in e-science and use of digital tools for research. The list is not to be considered as a "complete list of leading groups in data driven sciences", nor does it reflect "ranking" or assessment of "strategic importance". It does reflect a first set of inputs from one individual contact/partner university. As the ENLIGHT community develops and go further into AI and digital innovation, more examples will be added. *NB: Even though expertise may be associated with sustainable structures as "departments", projects will be terminated and new ones appear. Information in the table may change and is preferably seen as snapshot indicators of were "forerunners" can be found in different scientific areas within ENLIGHT.*

5) Relationships with the European Open Science Cloud[3]

Is your institution an EOSC member?
9 svar

● Yes
● No

44,4%

55,6%
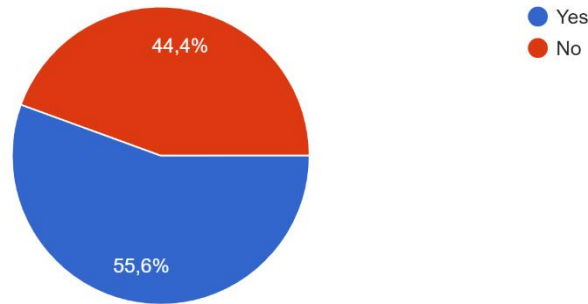
Five of the ENLIGHT partner universities are members of the European Open Science Cloud (EOSC). In addition, UPV/EHU are involved through the Basque Centre of Applied Mathematics. Relationships to EOSC are sometimes linked to national priorities and processes, thus affecting how an ENLIGHT university can act related to EOSC.

Not yet active: RUG, NUIG
Some activity: UPV/EHU, NUIG
EOSC Members: UU, UGOE, UT, UBx and UGent

Overall, the ENLIGHT alliance already has established a rather strong involvement in EOSC with more than 50 % of its partners active as EOSC members. UU, UGOE, UT, UBx and UGENT have internal groups for EOSC related activities and UU, UT, UGENT and UGOE are also involved in national coordination and EOSC Task Forces. It is a matter of definitions if ENLIGHT partners have strategies based on the EOSC Strategic Research and Innovation Agenda (SRIA)[4], since "yes" in some cases means that they are linked to national strategies and commitments. Estonia, Germany and The Netherlands have established national governance/coordination in this area, typically through the Mandated Organization (e.g. NFDI e.V. in Germany or BELNET in Belgium). Sweden has a kind of "soft" governance where national work is coordinated by the National Research Council (mandated organization) but association is voluntary.

Partners are also involved in infrastructures and initiatives connected to the EOSC, for example, ESCAPE (RUG), EOSC Nordic (UT, UU), EOSC Future (UGOE) and OpenAIRE (UGent, UT, UG

---

[3] https://eosc.eu/
[4] https://eosc.eu/sria

6) Involvement in ESFRI platforms

Answers indicated difficulties to assess the services and size of the user base of ESFRI platforms in which ENLIGHT partners are involved. More about ESFRI and platforms already up and running as well as planned ones are found via [5] ESFRI platforms are often ERICs with limited transparency to organizations. In cases where a university has persons in governance structures of a platform there is also information about users. At the European level some initiatives are still under development, e.g. ACTRIS, DiSSCo and SKA. However, some platforms are regularly used i.e. users who are familiar with processes and scientific environment:

1. ELIXIR/genomics: UU, UT, UGOE, UBx
2. PRACE and other HPC (high performance computing) initiatives: UU, UT
3. DARIAH (digital humanities): UGENT, UGOE, UU are active in digital humanities through a national platform.
4. LOFAR (astronomy): NUIG, RUG
5. ICOS (carbon dioxide observatory): UU active through a national platform
6. DiSSCo (biodiversity) – not yet operational but UU/Sweden and UT/Estonia are "observing" and ready to step up
7. EMPHASIS (plant phenotyping): UBx

7) Other areas of expertise

- Cyber security is added as an area of interest by ENLIGHT partners. The level of expertise was not mapped in this survey but there is reason to believe that it is possible to find persons to take part in or support activities where cyber security is crucial within ENLIGHT collaborative work. Competencies may be divided in two parts – support and research. Support is about to ensure that data and digital communication is secure. Most ENLIGHT partners also perform research related to cyber security.

- Augmented reality/VR, and cognitive AI was highlighted by RUG.

- AI and Societal transformation/impact is an area of expertise in UU (national doctoral school).

- Interest in upskilling activities was mentioned by several respondents. Research data management as an area of Open Science support activities is also targeted by WP 7 (cf. deliverable D7.1). From the perspective of ENLIGHT- RISE the approach on how to manage research data which is produced in the project is described in the projects Data Management Plan. To step up the awareness, a workshop on RDM for ENLIGHT

---

[5] https://www.esfri.eu/

participants will be organized in Fall 2022 (WP 7/"Open Science" is responsible for the workshop).

- Further comments in the survey were related to relationships between ENLIGHT E+ and ENLIGHT-RISE, e.g. support related to IT-systems for managing students and teaching activities used by the ENLIGHT platform. Such questions are not the scope of the present work package.

Sensitive data – Survey among UU researchers

The survey performed by input from all ENLIGHT partners demonstrates an institutional perspective and gives no information about how needs are expressed by individual researchers or research groups. One of the main challenges in data-driven sciences is how to share and do research on "sensitive data". According to the EOSC, sensitive data are defined as:

Data that needs to be protected against unauthorized access. Data protection may be required for legal or ethical reasons, for personal protection, or for proprietary considerations,

In addition, the scope of ENLIGHT-RISE is not only research but also innovation and to promote an ecosystem where DI/AI are core societal values. It is therefore also interesting to get some input on how users express their needs in this area.
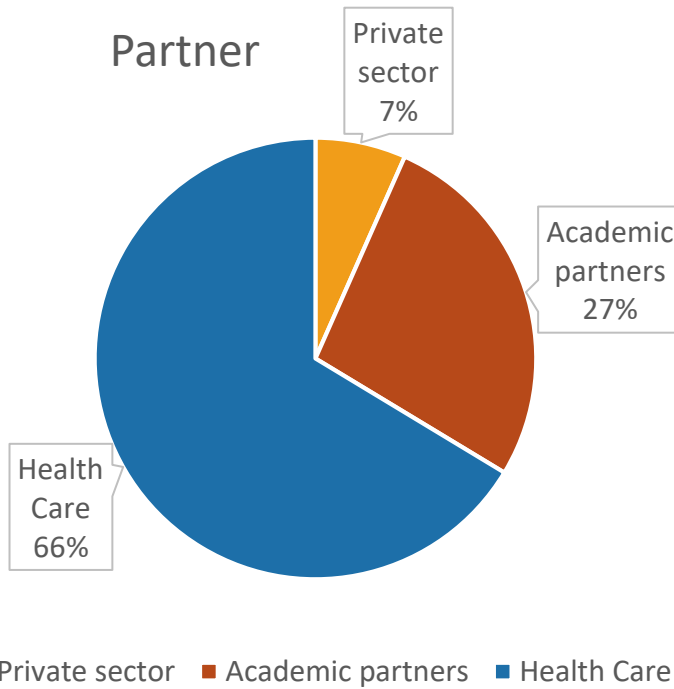
Uppsala university has performed an "in house" survey among researchers in order to get a first impression about challenges and needs associated with the interface between open and "protected" data. Since UU is a comprehensive university with nine faculties, a broad range of disciplines were targeted. The results are of course no complete overview describing the ENLIGHT community. However, it is an indication and thus useful for further discussions around how the needs can be described and taken care of in universities as well as collaborative projects with societal stakeholders. *N.B: It may be desirable to interoperate sensitive and non-sensitive data.*

A questionnaire was distributed to departments within the three disciplinary domains of Uppsala university - Humanities and Social Sciences, Medicine and Pharmacy and Science and Technology. 256 answers were collected. They represented a broad range of disciplines and people in rather senior positions/leaders of research groups or labs/services.

Results of survey related to the scope of ENLIGHT are presented below:

1) **Collaborating partners?**

Partner

- Private sector 7%
- Academic partners 27%
- Health Care 66%

■ Private sector ■ Academic partners ■ Health Care

2) **Data volume (estimated)**

Data Volume

Category	
> 100 TB/yr	
1-100 TB/yr	
100 GB-1 TB/yr	
1-100 GB/yr	
< 1 GB/yr	
Do not know	

0   20   40   60   80

ENL⩗GHT

This project has received funding from the European Union's
Horizon 2020 research and innovation programme under grant
agreement No 101035819
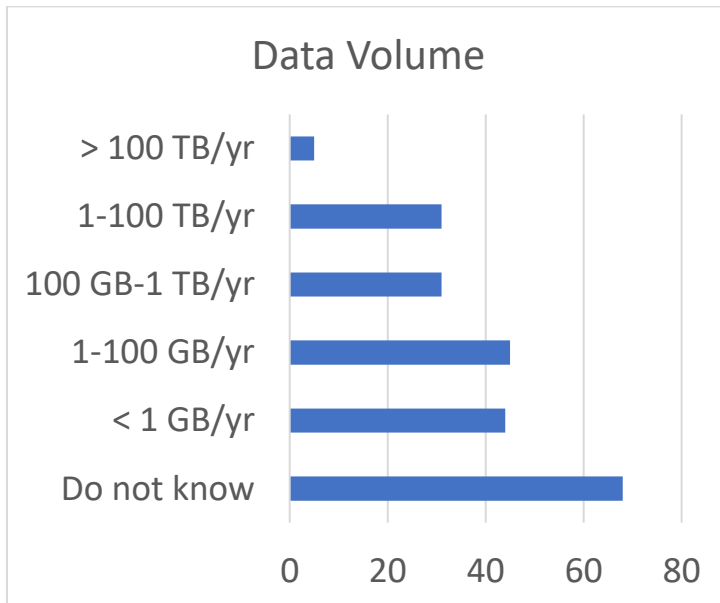
3) Need for support?
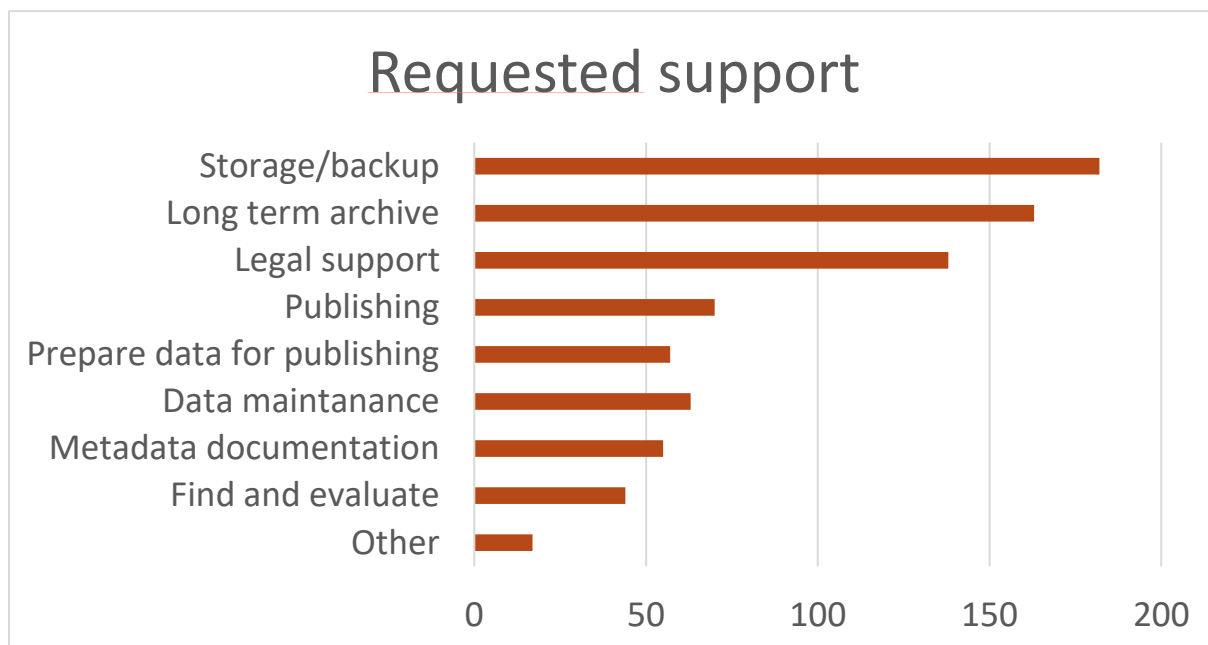
The survey indicates how needs are experienced by researchers using sensitive/protected data in their R&I projects. Collaborations involving the health care sector is a dominating source. Other kind of collaborations with business sector as well as strictly academic ones do also use and share "sensitive" data. Data volumes varies but it is noteworthy that many/about one third of the users do not know the size of their data sets or how much computing capacity they need. Resources for storage and archiving is a need expressed by many researchers but also request legal advice incl. IP for business/innovation activities. As the present work package further proceed this will be explored in e.g. case studies and dialogues with the Innovation network. The area of Open Science seems to be more mature (at least in Uppsala) due to the respondents.

DISCUSSION and CONCLUSIONS

This first mapping exercise among the ENLIGHT partners clearly demonstrates scientific strength and strategic commitment in digitalization and e-science. This is harmonizing with the European Commission setting out for a future Green, Digital and Open Europe.

Strengthening links between DRI resources and the ENLIGHT R&I agenda

The ENLIGHT partners already put large resources into digital platforms and services – and efforts are well in line with institutional scientific strengths and priorities. Excellence and commitment are also demonstrated when institutions and researchers involve and take part in governance and disciplinary networking in order to further develop capacity and competence. However, critical mass may be an indicator of commitment and long-term stability but it is not to be interpreted as scientific excellence. The nine ENLIGHT partners are different in many ways, with different strengths and opportunities – to be new in a field may be good enough.

1. A rough estimation of digital maturity and access to DRI-resources/services among the ENLIGHT partners
   a. Universities with a long-term commitment and broad activities on local, national and international level: UU, RUG, UT, UGOE, UGENT
   b. Universities with activities in some areas: UBx, NUIG
   c. Universities new in the DRI-field: UPV/EHU and UC

2. Correlation with mapping of existing collaborations is coherent and promising – indicating that infrastructural resources needed for the ENLIGHT collaborative work can be obtained within the ENLIGHT platform. Presence of DRI resources coincide with institutional strengths in research and education – and already existing collaborations. The finding also indicates that there is no need to create an "ENLIGHT digital research infrastructure" - connecting, sharing and being a part of existing efforts related to e.g. ESFRI science clusters and EOSC will be a more efficient model.

3. There was an attempt to map the identified DRI resources and expertise to the five flagship areas. It turned out to be impossible. The flagship areas are broad and multidisciplinary – while DRIs are of two kinds:
   - resources more or less related to disciplinary areas of research and education, or
   - generic resources, e.g. computing, AI and large amount of data to be interdisciplinary combined for new insights. It is in fact one of the major promises of e-science and open data that cross-disciplinary use of methods and data sets will open up for new ways to conduct research, innovation and shape the future society.

4. Even though there is an awareness of the importance of digital resources for future research and many platforms already in place, challenges to be further addressed were found. The most important are described below:

- In some cases, it turned out to be a gap between resources and expertise actually present in local departments and the information available among supporting officers in central positions. This is a challenge to be tackled within the HEI. It is probably also associated with a need for funding and support from decision making bodies. If a HEI is already operating major DRI platforms, there is probably an internal flow of information – if not, information is staying with individual researchers.

- The area of DRI as well as many EU or national projects demonstrating "best practise" is a jungle of acronyms and logotypes. Even if acronyms are explained it is difficult for non-experts to understand what's behind – and new ones appear frequently. It reduces transparency for non-experts and increase the need for support. The kind of competence needed is associated with offices supporting research and innovation.

- The small survey targeting the needs of individual researchers using sensitive data is calling for further work. Case studies will give some input but more information about the size and needs in terms of storage capacity, access to legal advice etc. must be obtained. More information about how to get access to data centre services etc. is also needed.

- The concepts of Open and Digital are expected to transform the way we do science and innovation. Expectations may be high, but the journey has just started in many disciplinary areas. To actually "interoperate" data in broad thematic projects will continue to be a significant challenge for a long time to come. Academia is rather open for "sharing" but societal stakeholders may be hugely regulated and have limited experience in international co-working projects. Data quality is another factor to be aware of. The ambitions of ENLIGHT are high in this area and the present WP is preparing for future work after the ENLIGHT-RISE project period, but the challenge is demanding!

Map of DRI involvement

On the basis of input from ENLIGHT partners, an aggregated map of involvement related to the European major initiatives in the interface of ESFRI/EOSC is presented below. More information about science clusters are found via [6][7][8][9][10]Involvement in major European initiatives is associated with local resources and expertise:
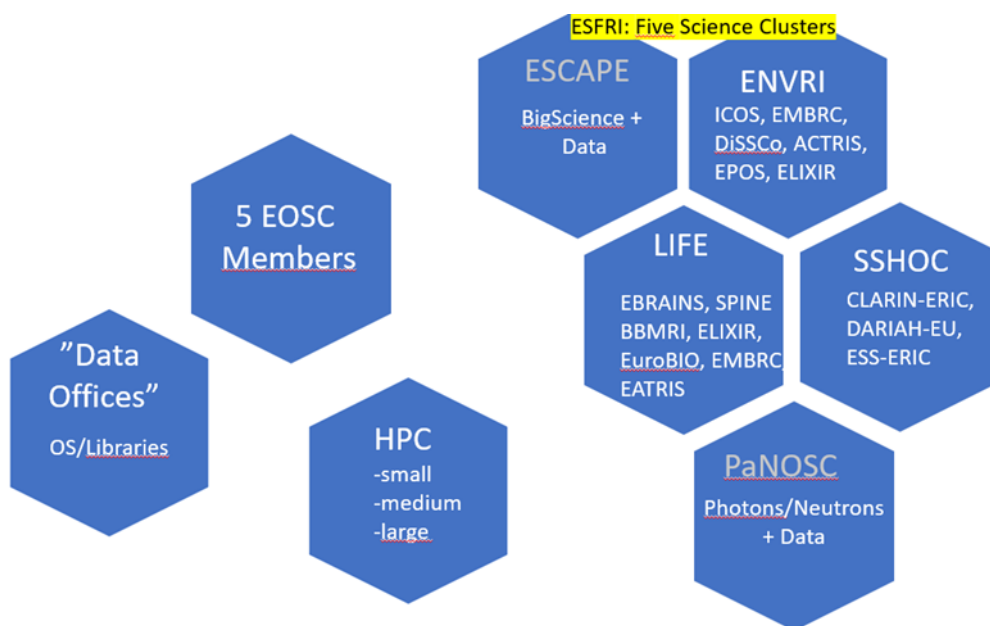
RECOMMENDATIONS & ENLIGHT ROADMAP to shared DRI

The results of the mapping exercise, as described above, open up for a more or less limitless number of possible collaborations. This is very much in line with the nature of research infrastructures as generic resources open for a wide spectrum of projects (platforms are usually based on "techniques" rather than scientific or societal "questions"). So are also the thematic areas of ENLIGHT with broad and multidisciplinary focus of relevance to societal challenges.

---

[6] https://projectescape.eu/

[7] https://envri.eu/

[8] https://www.eosc-life.eu/

[9] https://sshopencloud.eu/

[10] https://www.panosc.eu/

Method to connect and share/get access to DRI of relevance to ENLIGHT projects

The field of DRI is rapidly developing – and involves the use of new and already existing data to be further processed via e.g. AI and in the innovation system! In order to be prepared to take an active part in a future Digital Europe, the present work package is aiming at a situation where the ENLIGHT community has developed mechanisms and workflows to connect and share already existing DRI – resources thus fostering collaborations and synergies between ENLIGHT-partners for education, research and innovation.

Since the ENLIGHT community is well positioned to take part in ongoing European work in this area, a methodologic platform for future work within the ENLIGHT alliance based on guiding principles aligning with existing structures for research data life cycle/DRI is outlined below:

1) **Federated workflows within ENLIGHT** – The partner universities are responsible for their own internal processes, staffing and funding of *in house*-work. Output is then coordinated over the alliance by mechanisms approved by decision making bodies. Of special interest for the long term sustainability in the infrastructure area are the local presence of support for Data Management and offices for support to Research and Innovation in the ENLIGHT universities. How these are organized may vary between partners - however, the alliance partners are responsible for the capacity to provide information about local digital resources and competencies.

2) **Rely on accepted principles for connecting and sharing DRI** – Connecting and sharing DRI resources is at present an undertaking going on all over Europe. Work is organized on three levels: *Local, National* and *European*.

   The ENLIGHT alliance take part in existing workflows, governance models and planning with a long term scope of making first class digital resources available for ENLIGHT projects. Funding for project use must be considered outside this WP – resources are normally not provided for free. Within ENLIGHT, resources are continuously developing, e.g. the E+ Interconnected campus and a catalogue of best practise in different areas. Examples of "best practise" related to e-science and the use of digital infrastructure will be fed into the ENLIGHT Observatory.

3) **Plug in to ESFRI Science clusters and EOSC** – On overarching, European level, frameworks not only for connecting and sharing but also for developing the European Research Area/ERA is given priority and funding. The main initiatives related to DRI are ESFRI and EOSC. ESFRI organizes via five science clusters as outlined in the map above, with the enrolment of ENLIGHT partners visualized. Where ENLIGHT partners operate national nodes there are also local competence recognized by the community of scientific users. These are then interacting with EOSC as described in figure below – and so is the ENLIGHT alliance. The partnership EOSC will thus form a matrix coordinating connecting and sharing of digital resources all over Europe:

Interface to EOSC – in general: Think of ENLIGHT as a "plug-in"
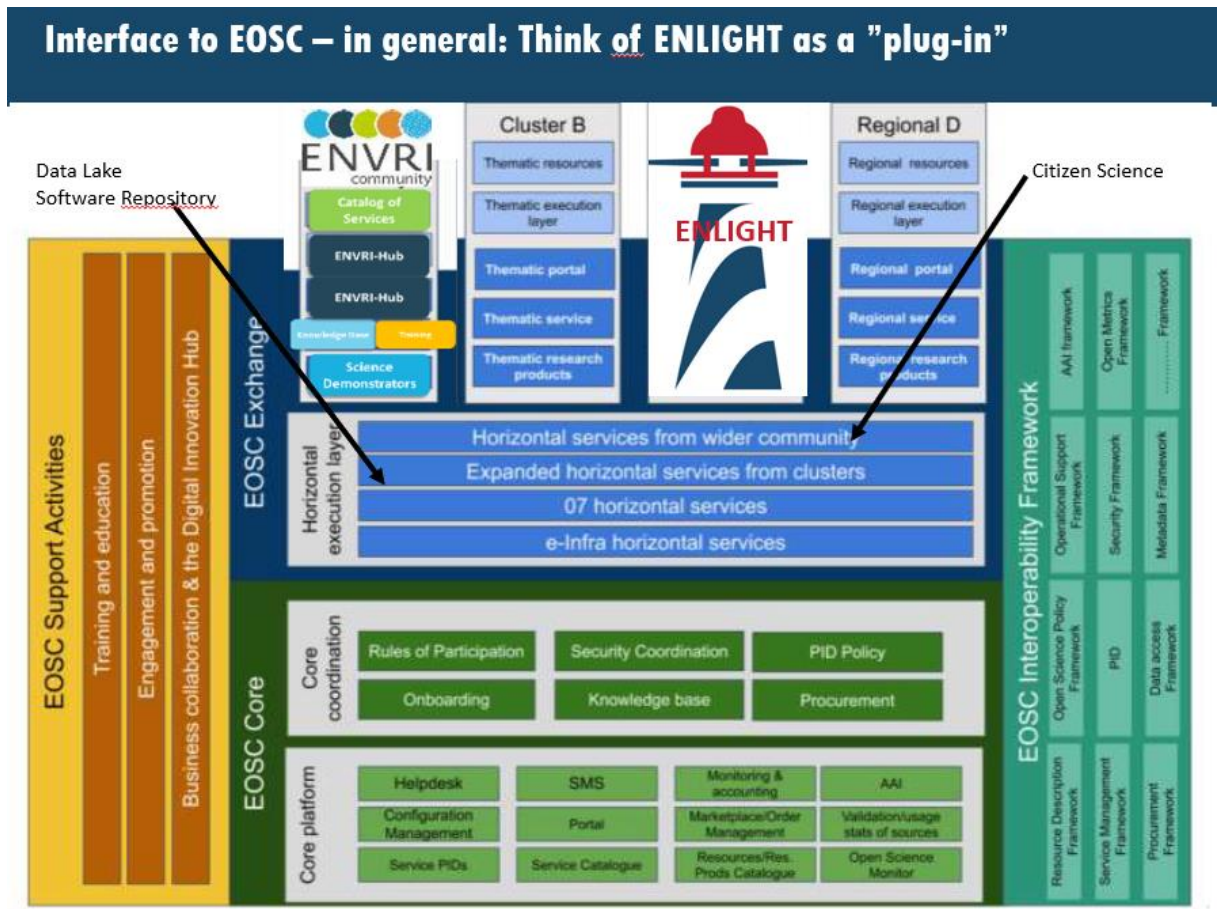
4) **Open up for innovation and projects with societal stakeholders and non-academic partners** - The ENLIGHT Alliance is building networks with society through Regional Academies, Think Tanks etc. There will also be a DI/AI Network to further explore DI/AI opportunities within the ENLIGHT Ecosystem with focus on Flagship Challenges. These networks are building blocks for the ENLIGHT platform and when projects are formed. On European level there are also work going on to set up Digital Innovation Hubs providing support and digital capacities for SME:s in different areas, e.g. health care. These are following the structure already outlined with local, national and European resources with relationships to ERUF.

At this stage, it must be stressed that differences in legislation and administrative/business traditions between countries are – and will be so for a long time – an ongoing struggle to overcome. EUN/ENLIGHT is a tool and an effort to take on the challenge but not a final solution to hand over a R/I-community where all obstacles are removed after a first three-year period.

5) **Upskilling** – In a rapidly developing area, access to upskilling is a core value. Upskilling activities and courses are already provided to some extent, often by e.g. national platforms with disciplinary scope or in doctoral schools. These are often openly announced and without fees. The ENLIGHT Alliance take part in existing initiatives but may organize internal upskilling activities when needed. A first example is a course in Data Management/Open Science for ENLIGHT Staff is offered by WP 7/Open Science in Q4, 2022.

## ROADMAP for sharing and connecting digital research infrastructure

A strategic roadmap is a visualized strategic plan to describe a desired future situation in terms of goals, timelines and intermediate steps. A Roadmap has the following functions:

1. Define present status – in this case promising and with many competitive resources but information as well as ENLIGHT workflows are still un-coordinated,
2. Outline a desired future "destination" and
3. Identify efforts and initiatives to get there.

The present status is outlined above including major principles for the upcoming work within the ENLIGHT Alliance (1). Even though many digital resources are available within ENLIGHT, and even more can be accessed through national and international initiatives, work is needed to further strengthen relationships and collaboration around DRI in the ENLIGHT community. Along with the project time line with Milestones and Deliverables, a digital ENLIGHT ecosystem need more of connection, networking and collaboration over organizational borders. In order to strengthen and deepen dimensions of interconnectivity five focus areas will be targeted in a cumulative workflow where each new step build on lessons learnt by the step before.

> *"The desired "destination" is an ENLIGHT Alliance where R&I projects in the five flagship areas can get access to digital research infrastructural resources and support of high international standard (2)."*
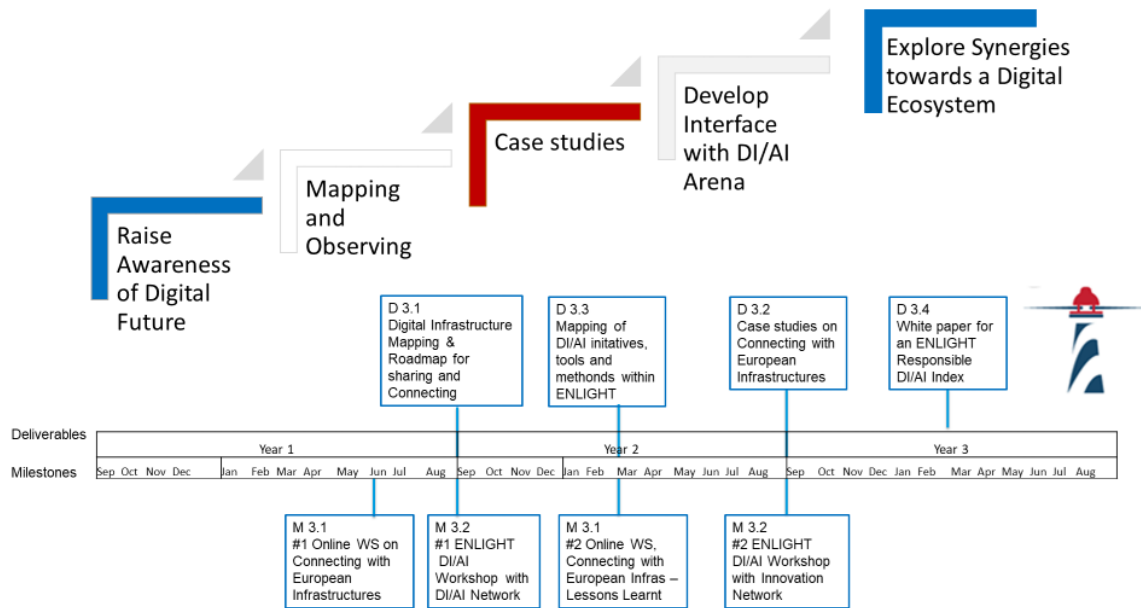
Raise Awareness – An ongoing focus is to raise awareness of the digital future in Europe. This focus area is continuously addressed in monthly WP-meetings, Project Board Meetings and ENLIGHT conferences. The federated work model encourages each ENLIGHT university to gain local competence in this area including an awareness of local profile/areas of competence. The WP organizes workshops and communicate in different ways, e.g. website, Youtube and reports, in order to put digital infrastructure on the ENLIGHT agenda.

Mapping and Observing – Gaining information about the present status of existing as well as important ongoing initiatives is an important roadmap effort. Through surveys within the ENLIGHT community, information about DRI platforms, open services and DI/AI initiatives are obtained. By involvement in existing major national and international efforts as e.g. EOSC and The Guild, the ENLIGHT community may connect with international frontiers of DRI, AI and policy making in this rapidly developing area. The latter is strongly encouraged by the Rectors and rely on a readiness to be open and keep up a dialogue with people in governance positions.

Case studies – On the basis of initial mapping activities, a set of three case studies will be performed. The cases are selected in order to further substantiate experience and practice around sharing and connecting with DRI. How to actually integrate/interlink will be further explored. Cases will be related to some of the ESFRI Science clusters and ENLIGHT Flagship Challenges.

Develop Interface with DI/AI Arena – In parallel with the case studies, the mapping efforts will deliver further input about existing AI initiatives, tools and methods within ENLIGHT. A DI/AI Network and connections with Innovation Network will also be established. Workshops will be arranged in order to develop the digital interfaces with AI and also explore beneficial aspects of innovation in the digital area in collaboration with WP 5.

Explore Synergies towards a Digital Ecosystem – Several initiatives of relevance are developing within the ENLIGHT-RISE project. Connections and opportunities related to deliverables in other work packages are important as the work proceeds. Special attention will be on Open Science with direct connections outlined in the Research Data Life Cycle (WP 7) and the role of Open Science Ambassadors. There are also initiatives related to early career researchers - an important group where digital skills will be essential for a future career in research (WP 4). Information about best practise in data driven research will continuously be shared by the Observatory (WP 2). The Interconnected Campus of the Erasmus+-project is another important interface to explore.

Ethical issues will now be considered and information/practices are aggregated in a White paper for a responsible DI/AI Index. The Index will provide guidelines to benchmark and maximize the positive impact of DI/AI on ENLIGHT Flagship Challenges.

Appendix – Examples of "best practise" in data driven science related to ENLIGHT Flagship Challenges

Question 4 in the Mapping Exercise was aiming at information not only about DRI but also projects/research groups demonstrating best practice as users of DRI and data driven methods in research. In the table below each ENLIGHT partner is presented by two rows:

- The upper row presents information about digital infrastructures related to the university. The information can be interpreted as computing capacity/databases and personnel for support to users, etc. Strong connections to national/international resources normally build on the presence of local resources.

- In the row below there is information in acronyms about projects etc. demonstrating best practice and *know-how* as advanced users of data and digital resources for research. More information about these examples/projects/groups will be fed into the ENLIGHT Observatory by UU.

The information below may seem "volatile" and as a first shot without a firm basis of definitions. The question was also "open" in the survey – thus interpreted in different ways by respondents and followed by iterative dialogues in some cases. This is an area were centrally located officers often have limited information about what is actually present on the "research floor". However, when ENLIGHT is going into an active project phase there must be researchers with experience present to successfully take responsibility for and plan the projects There will hardly be a situation where there is an ENLIGHT "project management hub" responsible for making things happen. There must be either an experienced PI, a societal partner with resources/experience or both (preferably) to successfully apply for funding etc. More information about experience and resources in partner organizations outside the ENLIGHT academic partners will be obtained in connection with work forming the DI/AI network and interaction with Innovation District. The infrastructural platforms themselves are also in many cases an arena for collaboration with contacts/users from outside academia. In fact, infrastructural platforms share many similarities with the ENLIGHT platform, aiming not only for excellent research but also collaborations outside the HE sector and societal benefits.

	Health and well-being	Digital revolution, e-science and AI in general	Climate change and environment	Energy and circular economy	HSS	Generic resources for computing and data management
UBx	SPINE Bioimaging, ELIXIR	TRANS-MATH		AENIGMAE Ecodesign, Green Concrete,		The Hub, Computing IS,
	AI-chair; neuroprocessing and non invasive detection	AI chair; signal processing (neutrino dect)		AI-chair; green AI		
UGent	ELIXIR		EMBRC		DARIAH, Gent C. Digital Humanities,	HPCGent Flemish Supercomputer Center. EuroCC,
UC						HPC/CLARA
	Computational pangenomics			Centre for materials science – in progress		
UPV/EHU		Basque C. Applied Math				.
NUIG	C. Clinical Data Management,	Multi-Level SmartCity,	ICOS, AEROSOURCE, ACTRIS,			Insight Data Analytics/Ireland, ICHEC/HPC,
UGOE			NFDI4Biodiversity, NFDI4Earth, NFDI4Chem		DARIAH, TextGRID	
RUG	Data Science in Health, Biobanks, MOLGENIS	AI Lab/Center, Spatial Data/GIS, 5FieldsLab/IoT, VR Environment, Reality Center			Euclid network, EESSI	HPCI, OpenStack Platform, RDMS,
	Covid-19		Space-related projects			
UT	BBMRI, ELIXIR, Estonian centre of Genomics		DiSSCo		SoBigData++, ESS ERIC, CLARIN ERIC	HPC/ETAIS/NeIC,
	Centre for Clinical research	Robotics Estonia - manufacturing				
UU	SciLifeLabDataCenter, Biobanks, National hub for sensitive data/bioinformatics, BBMRI, EATRIS	AI4Research, WASP-HS, Spatial Data/GIS,	EMBRC, DiSSCo (pending), ICOS, ACTRIS,		CLARIN-ERIC, DARIAH, peace and conflict database, DigiArk,	HPC/UPPMAX, EuroCC, NeIC,
	Covid-19, U-CARE, U-CAN, Centre for clinical research	SIP-IoT	Four CC related to Food	InnoEnergy, Battery2030, EIT RAW, Materials 2.0	INCULTUM, EIT Culture and Creativity	