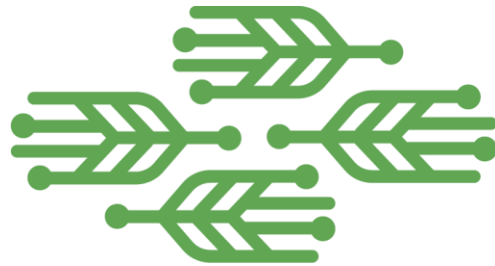


Deliverable 3.2 – WP3. Report on HIKR design



EURAKNOS

***Connecting Thematic Networks as Knowledge Reservoirs towards a European
Agricultural Knowledge Innovation Open Source System***

Task 3.2 Database, technology choices & e-KRP recommendations

Summary

The overall aim of EURAKNOS is to connect all the materials, knowledge and platforms of existing TNs EURAKNOS in one electronic Knowledge Reservoir Platform (e-KRP) to map the stored knowledge within each network, and to design a common data system to make this knowledge better accessible, findable, interoperable and reusable for the agricultural innovation community in Europe. Deliverable 3.2 describes the different steps and conclusions made in order to define the design and tools (software, data formats) to be used for a HIKR in perspective of construction of the prototype of the common e-Knowledge Platform (e-KRP), and communication and dissemination perspectives to the end user. In order to achieve this goal a workshop with experts was held on different topics related to creating the ontology and building the database. This allowed us to describe the EURAKNOS data model and the technological framework (including the data storage, metadata storage and the search engine). To conclude we present a strategy to move from the HIKR to the e-KRP, including the recommendations on how to attain a user-friendly and accessible digital platform.

| Deliverable Number | Work Package |
|---------------------------------------|---|
| 3.2 | WP3 |
| Lead Beneficiary | Deliverable Author(S) |
| LFG | Davino Van Hal, Laurens Van der Cruyssen |
| Beneficiaries | Deliverable Co-Author (S) |
| IDELE, PSKW, ARC, NAK, RAU, UGent, AU | Hercules Panoutsopoulos, Sofia Mouseti Borja Espejo Garcia |
| Planned Delivery Date | Actual Delivery Date |
| 30/04/2020 | 30/04/2020 |

| | | | |
|---------------------|-----|---|---|
| Type of Deliverable | R | Document, report (excluding periodic and final reports) | X |
| | DEC | Websites, patents filing, press & media actions, videos | |
| | E | Ethics | |

| | | | |
|---------------------|----|--|---|
| Dissemination Level | PU | Public | X |
| | CO | Confidential, only for members of the consortium | |



Content

| | |
|--|-----------|
| List of figures | 4 |
| List of tables | 4 |
| List of abbreviations | 5 |
| 1. Introduction | 6 |
| 2. Methodology..... | 6 |
| 3. Design WG Meeting..... | 8 |
| 3.1 Working Group objective | 8 |
| 3.2 Working Group participants..... | 8 |
| 3.3 Working Group content & outcome | 8 |
| 4. Design of EURAKNOS HIGH IMPACT KNOWLEDGE RESERVOIR | 10 |
| 4.1 Framing the concept of “High Impact Knowledge Reservoir” | 10 |
| 4.2 EURAKNOS data model | 11 |
| 4.2.1 What is a data model, and why is it needed? | 11 |
| 4.2.2 How can Thematic Networks work with the EURAKNOS HIKR? | 12 |
| 4.2.3 Towards an ontology for formally describing digital outputs of Thematic Networks | 13 |
| 4.2.4 Presentation of the EURAKNOS data model | 15 |
| 4.2.5 Compliance with the FAIR principles | 17 |
| 4.3 Considerations for selecting data objects to store into the High Impact Knowledge Reservoir | 18 |
| 4.4 Technology framework..... | 23 |
| 4.4.1 Data storage..... | 24 |
| 4.4.2 Metadata storage | 25 |
| 4.4.3 Search engine..... | 25 |
| 5. From HIKR to e-KRP (Digital Platform) | 28 |
| 5.1 Types of knowledge reservoirs in current TNs | 28 |
| 5.2 A (preliminary) future vision for the e-KRP in HIKR | 28 |
| 5.2.1 Phase 1: Prototype - LIBRARY | 29 |
| 5.2.2 Phase 2: Minimum Viable Product (MVP) – Platform..... | 30 |
| 5.2.3 Phase 3: Farmbook | 30 |
| 5.2.4 Future | 31 |
| 5.3 General usability and accessibility guidelines for the e-KRP | 31 |
| 5.3.1 Recommendations on usability & accessibility..... | 32 |
| 6. Conclusions & Next Steps | 33 |
| 7. Annex 1: Classes of the EURAKNOS ontology and description of their properties | 34 |
| 8. Annex 2: Details about the advisor survey conducted in the context of the EUFRAS meeting in Athens | 44 |
| 9. Annex 3: State-of-the Art in Usability & Accessibility Guidelines | 48 |
| 9.1 Usability guidelines | 48 |
| 9.1.1 Accessibility principles and guidelines | 52 |
| 10. References | 54 |

List of figures

| | |
|---|----|
| Figure 1: Task 3.2 Methodology overview | 7 |
| Figure 2: the basic pillars of the EURAKNOS HIKR | 10 |
| Figure 3: Scenarios illustrating the potential contribution of the EURAKNOS data model and repository to future Thematic Network efforts | 13 |
| Figure 4: Methodology adopted for the work undertaken towards the development of the EURAKNOS data model | 13 |
| Figure 5: The EURAKNOS data model | 16 |
| Figure 6: Process from the identification of the e-KRP end-user types to the integration of TN data into the EURAKNOS HIKR | 18 |
| Figure 7: Advisor responses to the question “What kind of content/information would you prefer?” | 20 |
| Figure 8: Advisor responses to the question “Which of the following formats of digital content is most preferable to you?” | 21 |
| Figure 9: Kinds of Thematic Network outputs to be integrated into the repository of the prototype e-KRP; the box heights are indicative of number of items per output kind to be considered in relation to the other kinds..... | 23 |
| Figure 10: Basic architecture of the Solr search engine | 27 |
| Figure 11: Schema.org types, in the context of the hierarchical structure employed by the schema.org vocabulary, used for identifying and defining properties for the EURAKNOS ontology’s classes | 35 |
| Figure 12: Gender distribution in the sample of advisors that participated in the EUFRAS meeting survey | 44 |
| Figure 13: Distribution of ages in the sample of advisors that participated in the EUFRAS meeting survey | 45 |
| Figure 14: Nationalities of survey participants | 45 |

List of tables

| | |
|---|----|
| Table 1: Potential ontology entities, entity properties and relations between entities | 15 |
| Table 2: Alignment between end-user types and their needs in terms of Agriculture-related knowledge (adapted from Gernert et al., 2019) | 19 |
| Table 3: Alignment between advisors’ preferences and the most relevant kinds and types of outputs proposed in the data model | 21 |
| Table 4: Alignment between classes of the EURAKNOS ontology (data model) and types of the schema.org vocabulary | 34 |
| Table 5: Properties of the <code>ThematicNetwork</code> class of the EURAKNOS ontology | 35 |
| Table 6: Properties of the <code>Domain</code> class of the EURAKNOS ontology | 36 |
| Table 7: Properties of the <code>Topic</code> class of the EURAKNOS ontology | 36 |
| Table 8: Properties of the <code>Output</code> class of the EURAKNOS ontology | 37 |
| Table 9: Properties of the <code>EndUserType</code> class of the EURAKNOS ontology | 39 |
| Table 10: Properties of the <code>Creator</code> class of the EURAKNOS ontology | 39 |
| Table 11: Properties of the <code>Audio</code> class of the EURAKNOS ontology | 39 |
| Table 12: Properties of the <code>Dataset</code> class of the EURAKNOS ontology | 40 |
| Table 13: Properties of the <code>Document</code> class of the EURAKNOS ontology | 40 |
| Table 14: Properties of the <code>Image</code> class of the EURAKNOS ontology | 41 |
| Table 15: Properties of the <code>Presentation</code> class of the EURAKNOS ontology | 41 |
| Table 16: Properties of the <code>Software</code> class of the EURAKNOS ontology | 41 |
| Table 17: Properties of the <code>Video</code> class of the EURAKNOS ontology | 42 |



| | |
|--|----|
| Table 18: Properties of the <i>Individual</i> class of the EURAKNOS ontology | 43 |
| Table 19: Properties of the <i>Organisation</i> class of the EURAKNOS ontology | 43 |

List of abbreviations

API: Application Programming Interface
 BSD: Berkeley Software Distribution
 CMS: Content Management System
 EIP-AGRI: European Innovation Partnership for Agriculture
 e-KRP: electronic Knowledge Reservoir Platform
 EUFRAS: European Forum for Agricultural and Rural Advisory Services
 EURAKNOS: Connecting Thematic Networks as Knowledge Reservoirs: towards a European Agricultural Knowledge Innovation Open Source System
 EUREKA: European Knowledge Repository for Best Agricultural Practices
 FAIR: Findable Accessible Interoperable and Reusable
 HIKR: High Impact Knowledge Reservoir
 KIP: Knowledge Innovation Panel
 KR: Knowledge Reservoir
 MA: Multi-Actor
 MS: Member State
 MVP: Minimum Viable Product
 PSKW: Proefstation voor de Groenteteelt Sint-Katelijne-Waver
 RDF: Resource Description Framework
 SQL: Structured Query Language
 TN: Thematic Network
 UML: Unified Modelling Language
 URL: Uniform Resource Locator
 WG: Working Group
 WHO: World Health Organisation
 WP: Work Package
 XML: Extensible Markup Language

1. Introduction

A common challenge for Thematic Networks (TNs) is the sustainability of the knowledge collected after the funding period of the project. Given the needs for financial resources to store TN outputs, ensure easy access, and regular updates in the long term to the technical infrastructure, a substantial part of the outputs of TNs is at risk of becoming easily outdated or unavailable to relevant end users such as advisory services, farmers and foresters.

Moreover, TNs are very diverse in their topics and scope and produce a wide variety of dissemination material, making it hard for end users to maintain a clear overview of the knowledge available that could provide potential solutions to the problems they seek guidance for. Additionally, given that information is scattered over a multitude of platforms & websites, end users make use of popular search engines such as Google, Bing or Ecosia in order to find the knowledge they need. However, the provided results are not always accessible (e.g. paywall), available in the user's native language, easily applicable to their context, or trustworthy.

Finally, creating a digital knowledge reservoir can be a daunting challenge for TNs. The end product may be a platform that indeed makes knowledge available, but lacks a consistent design and/or violates general principles of usability towards end users (e.g. a platform that enables no easy access to data). These challenges are the scope of Task 3.2: Design of a HIKR (High-Impact Knowledge Reservoir).

The main goal of Task 3.2 is to identify in a participatory way the best design and format for a High Impact Knowledge Reservoir (HIKR) focusing on farmers, foresters & advisors. Task 3.2 will carry out activities to define the design as a common format and tools (software, data formats) to be used for a HIKR in perspective of construction of the prototype of the common e-Knowledge Reservoir Platform (e-KRP) (WP4) and communication and dissemination perspectives to the end user.

Within this context, the aim of this document is to frame the concept of the High Impact Knowledge Reservoir (HIKR), present and describe the EURAKNOS data model (core to the design of the HIKR), based on the methodology presented next, illustrate a number of considerations with regard to the types of data (and formats of the content and information conveyed) needed to be stored into the HIKR, as well as present a future vision for the EURAKNOS electronic Knowledge Reservoir Platform (e-KRP) through the HIKR.

2. Methodology

Task 3.2 works on the outputs of Work Package 2 (D2.2) and of Work Package 3 (D3.1).

First, preliminary work was performed on the definition of the concept of Thematic Networks and the identification of key terms & the relations between them. This early work was the foundation for the development of the draft version of a potential EURAKNOS data model.

Subsequently, a 2-day workshop (WP3) was organised in Paris on 12th and 13th December 2019. The input for this workshop came from exercises and discussions during the Athens consortium meeting (WP1), as well as the Budapest workshop (D2.5), and aimed at identifying and making decisions with regard to the scope of the HIKR, the EURAKNOS data model, the features required, and the future vision of the HIKR for TNs and potential end users.



D3.2 – Report on HIKR design

To further elaborate on the outcomes of the working group, Task 3.2 held weekly sessions with a team of 10 technical experts, suggested by the KIP (Knowledge Innovation Panel) and EURAKNOS consortium for their technical background and experience in complex IT projects, to discuss the topics of data model and technology-related choices. Based on these discussions, a technology stack and a phasing of the EURAKNOS platform - beyond the EURAKNOS project - was suggested.

Furthermore, a survey to advisors was conducted during the EUFRAS meeting¹ in Athens. The survey took place on the 25th of February 2020 and was implemented by administering questionnaires to a sample of 32 advisors.

Next, the state-of-the-art of accessibility & usability guidelines was researched as input for the prototype & future vision which will be developed in Work Package 4.

The insights, conclusions & recommendations are compiled in this deliverable D3.2.

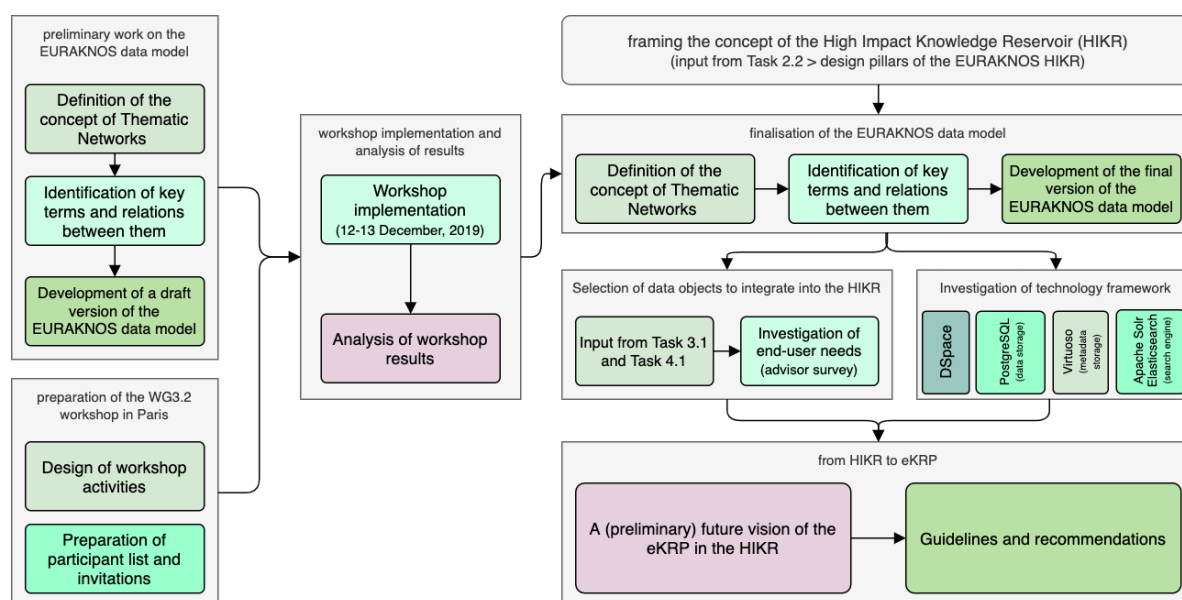


Figure 1: Task 3.2 Methodology overview

¹ <http://www.eufra.eu/index.php/activities/archive/142-eufra-annual-assembly-2020>

3. Design WG Meeting

3.1 Working Group objective

The DoA states that "The design WG will be set up to identify and file recommendations for best methodologies and practices to design a HIKR. Design WG members (ca. 10) will be selected from the KIP. The findings will be synthesized in a report discussed and validated by all the partners involved in the work package, describing the methodologies and practices to design a HIKR."

3.2 Working Group participants

The workshop participants were suggested by the KIP (Knowledge Innovation Panel) and EURAKNOS consortium. During this phase in the project, Task 3.2 is mostly focusing on solving the technical challenges in setting up a High Impact Knowledge Reservoir (HIKR).

For this reason, only people with a highly technical background and experience in complex IT projects were invited to the working group. An invitation to the workshop was sent to a list of 30 experts.

The workshop was attended by a total number of twelve (12) experts coming from diverse backgrounds (research, IT project management, software and data engineering, digital design) and countries (Belgium, Denmark, France, Greece, Hungary, The Netherlands).

3.3 Working Group content & outcome

The goal of the WG3.2 workshop (i.e. the workshop related to the activities involved in the design of the HIKR), which took place in the EURAKNOS Paris meeting, was to present the initial design-related considerations to the guest participants and get involved in a fruitful brainstorming and discussion with the aim to receive feedback and make decisions about a number of design- and development-related issues.

The overarching research questions on which the WG meeting focused were the following:

- How can data available from TNs be collected and integrated into a HIKR?
- In which ways could the EURAKNOS data model serve as a basis for future TNs?

The above research questions were further analysed in a number of topics that were used as a guide for the discussion implemented in the context of the workshop:

1. Technical issues that need to be tackled.
2. Scope and expectations from WG3.2 working group.
3. EURAKNOS data model: Can it adequately and efficiently represent the structure of data? Do we need to further elaborate and extend upon it?
4. Need to comply with Open data standards and FAIR data principles.
5. Data integration.
6. Search engine technologies to consider.
7. HIKR interoperability and implications for backend design.
8. End-user types, content needs and implications for search patterns.
9. Learning from TN best practices.

For each of the topics (topics 1 to 5 were those that were more extensively addressed), an introduction and initial, design-related considerations were presented to the group of technical experts. Topic per

topic feedback was gathered from every participant in the working group. The feedback received from the presentation of a draft version² of the EURAKNOS data model³ indicated that the data structure proposed is adequate for providing a solution to the problem at hand (i.e. the adequate and efficient storage of the range of outputs available by the KRs of TNs).

A significant contribution of the workshop related also to the recommendations received with regard to the technologies that should be considered for developing the HIKR⁴ (and, thus, the e-KRP⁵). More specifically, one of the topics addressed in the WG3.2 workshop related to the need to “*comply with Open data standards and FAIR data principles*” and the implications for the selection of technologies. The solution that had been initially considered for the e-KRP, and presented to the workshop, was based on a stack of technologies comprising a popular NoSQL⁶ database system (i.e. MongoDB) and a search engine system built on top of it (based on either Solr or Elasticsearch). According to the feedback received, such a solution could indeed fit the purpose of EURAKNOS and work, but, given the fact that it is a custom-made, in-house solution, there could be difficulties in its maintenance especially after the end of the project. It was suggested that attention should be paid to existing frameworks for creating open repositories able to not only deliver content, but also to manage and curate it. Such solutions allow to easily comply with the Open data standards and FAIR principles, which is a requisite for the EURAKNOS HIKR and e-KRP.

In this context, DSpace has been proposed as a solution interesting to investigate and adopt, given its rich set of features, the vibrant community built around it and the number of projects that have been built on it. Based on the feedback received: (i) the conceptual design process (which is documented in the following chapters) has been continued and resulted in the data model presented in this document; and (ii) the process of the HIKR (and e-KRP) development, which is currently in progress, has been initiated.

² The final version of the EURAKNOS data model described in this document is based, in almost its entirety, in the draft version presented in the WG3.2 workshop in Paris. It has been produced as the result of fine tuning the draft version with the aim to better capture and illustrate important characteristics of the digital outputs (i.e. data objects) available by TNs. In order to keep the descriptions provided clear and understandable, the final outcome of the design process (i.e. the final version of the EURAKNOS data model) is presented.

³ A definition of the term “data model” together with an explanation of its role and need are provided in the sub-section 4.2.1 (“What is a data model and why is it needed?”).

⁴ The concept of the HIKR is extensively addressed in the “Framing the concept of ‘High Impact Knowledge Reservoir’” section of the document. It is used to refer to the data repository lying at the back-end of the EURAKNOS electronic Knowledge Reservoir Platform (eKRP).

⁵ The term “e-KRP” is used to denote the EURAKNOS electronic Knowledge Reservoir Platform in its entirety including both back-end (i.e. the HIKR) and front-end components (i.e. the user interface).

⁶ According to Wikipedia (“<https://en.wikipedia.org/wiki/NoSQL>”), NoSQL data stores provide data storage and retrieval mechanisms other “*modeled in means other than the tabular relations used in relational databases*”

4. Design of EURAKNOS HIGH IMPACT KNOWLEDGE RESERVOIR

4.1 Framing the concept of “High Impact Knowledge Reservoir”

TNs are multi-actor efforts oriented towards the collection and dissemination of knowledge (data) objects that convey agriculture and forestry-related content and information. As Marois et al. (2019) point out, data disseminated by TNs has been assembled into fit-for-purpose collections, termed as KRs, with the aim to facilitate the community’s access to knowledge. A KR is a collection of best materials, practices, instruments, methodologies, and tools that contribute to the use of innovative solutions for sustainable agriculture and forestry. From a technical perspective, a KR relates to the technologies and tools employed for facilitating data storage and end users’ access to content and information.

As made evident by the analysis of Marois et al. (2019), there is a number of metaphors that have been employed by the existing TNs for the needs of making agriculture-related content and information available to end users. These metaphors are:

- **KR as a catalogue** with an emphasis on the provision of functionalities for assisting and supporting the user in the search and retrieval of digital content.
- **KR as a library** with an emphasis on the classification of digital content with regard to the types of data objects available and the topics addressed.
- **KR as a practical, digital guidebook** with an emphasis on the provision of ready-to-use and easy-to-apply solutions tailored to the needs of the end users targeted.
- **KR as a marketplace (for the exchange of good practices)** with an emphasis on redirecting the user to web-based information environments of interest.

The diversity in the metaphors adopted has led to various solutions for the architecture of information and the structuring of data provided, which have been based on tailor-made, fit-for-purpose designs (Marois et al., 2019). As far as access to data is concerned, there are two approaches adopted: (i) data is available from the TN’s website; and (ii) data is stored and accessible from dedicated databases.

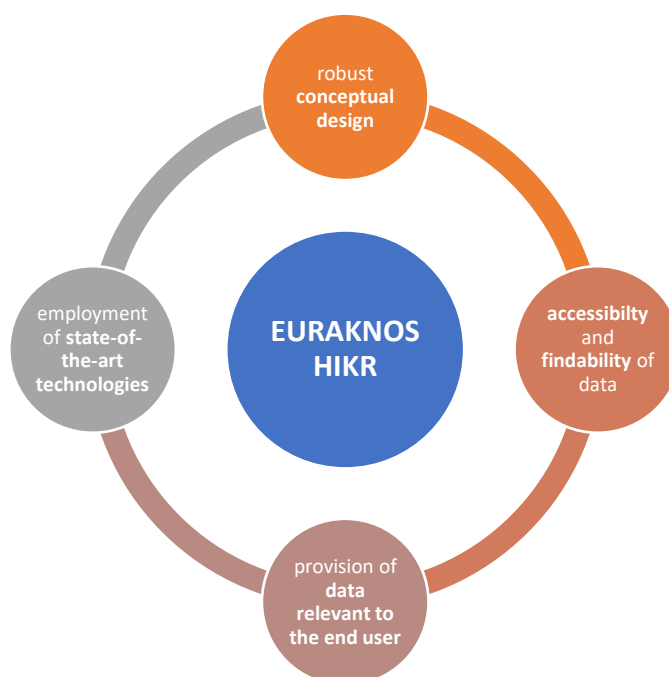


Figure 2: the basic pillars of the EURAKNOS HIKR

EURAKNOS aspires to create and provide a **High Impact Knowledge Reservoir (HIKR)**; in other words, a collection of data from TNs with an increased importance to the end users targeted (i.e. farmers, foresters, and advisors). Apart from being relevant, data should also be easily accessible and findable (Marois et al., 2019). Therefore, the development of the HIKR of EURAKNOS needs to be based on technologies that are compliant with open data standards and have the capacity to facilitate access by both human and machine agents. However, accessibility and findability are two core features that need to be built on a robust conceptual design. So, the basic pillars of the EURAKNOS HIKR, as shown in Figure 2 above, are:

- a robust conceptual design (i.e. the EURAKNOS data model);
- the accessibility and findability of the data provided;
- the availability of data that is important and useful to the end user types targeted; and
- the employment of state-of-the-art technologies that comply with open data standards and are able to facilitate access to data by both human and machine agents.

Based on the above, the EURAKNOS HIKR is going to have a state-of-the-art data repository at its core, where data will be stored according to the data model presented in the following section. It needs to be stressed that the term “data repository” holds a different meaning than that of “database”, despite the fact that these terms are sometimes used interchangeably. A database is “*an organised collection of data, generally stored and accessed electronically from a computer system*”⁷, whereas the term “data repository” relates to technologies of a broader scope by emphasising on data management and data curation⁸ issues. More specifically, a data repository can be conceptualised as a database-related infrastructure providing solutions for data collection and management with regard to a wide spectrum of purposes⁹. By taking account of the basic pillars on which the EURAKNOS HIKR needs to be built, as well as the feedback received at the Paris meeting’s workshop, the technology solution to be employed for the development of the EURAKNOS data repository, on the basis of the library metaphor (phase 1 of the EURAKNOS Platform) as stated in the first phase of the evolutionary path towards a future, fully-fledged digital platform (see section 7), is **DSpace**¹⁰.

4.2 EURAKNOS data model

4.2.1 What is a data model, and why is it needed?

The data objects that have been collected and disseminated by the existing TNs, which are intended to be integrated into the EURAKNOS repository, are heterogeneous in both their format and types. Apart from that, there is limited to zero evidence with regard to the data models employed by TNs. Therefore, the provision of content and information that is relevant to the end users targeted needs to be based on a robust design conveying the way in which data is modelled and structured in EURAKNOS. More specifically, a data model is “*an abstract model that organises elements of data and standardises how they relate to one another and to the properties of real-world entities.*”¹¹ So, given that “**output**” (i.e. any data object available by a Thematic Network’s Knowledge Repository) is the entity that is core to the EURAKNOS-related design, the goal of the EURAKNOS data model is to **capture** and **formally describe** all the **properties** that this entity possesses together with other entities relating to “output”, their properties, and any potential relations between these entities.

⁷ <https://en.wikipedia.org/wiki/Database>

⁸ According to Wikipedia (https://en.wikipedia.org/wiki/Data_curation), **data curation** is “*the organisation and integration of data collected from various sources*” and involves processes of “*annotation, publication and presentation of the data such that the value of the data is maintained over time, and the data remains available for reuse and preservation.*”

⁹ <https://digitalguardian.com/blog/what-data-repository>

¹⁰ More details are made available in the “**Technology framework**” section.

¹¹ https://en.wikipedia.org/wiki/Data_model

As stated above, there is great heterogeneity in the data objects provided by the KRs of existing TNs. However, in order to store and provide access to this data in an efficient way, there needs to be an adequate description of the different **kinds** of outputs (i.e. audio, dataset, document, image, presentation, software, and video) and the respective **formats** in which content and information is conveyed, the **types** of the data objects available, and the **purpose** that a data object is supposed to serve. These are **specific properties** of the “output” entity, described in the EURAKNOS data model, which have been modelled in a way that not only considers the existing Thematic Network landscape but also attempts to cover future efforts. “**Thematic Network**”, “**Domain**” (i.e. the area of interest that a Thematic Network is particularly related to), “**Topic**” (i.e. the theme that an output of a Thematic Network is about), “**Creator**” (distinguished into “**Individual**” and “**Organisation**”), and “**End User Type**” are entities that have been identified as necessary for an adequate and efficient description of the digital outputs available by the Knowledge Reservoirs of Thematic Networks and are presented in detail in the following sections.

4.2.2 How can Thematic Networks work with the EURAKNOS HIKR?

By providing a well-structured, formal description of the entities, entity properties, and relations that are core for an efficient description of Thematic Networks’ digital outputs, the EURAKNOS data model is considered to be a significant contribution to the Thematic Networks’ community. There are three scenarios (see Figure 3 below) for illustrating the potential contribution of the EURAKNOS data model (and repository) to future Thematic Network efforts:

- **Scenario #1:** A Thematic Network can *consume data available from the EURAKNOS repository* through exploitation of the Application Programming Interface (API)¹² presented and described in Deliverable 4.2 (i.e. “**API documentation**”).
- **Scenario #2:** A Thematic Network can have its own KR, but its data is also stored into the EURAKNOS HIKR by making use of the EURAKNOS data model.
- **Scenario #3:** A Thematic Network can *develop and maintain its own Knowledge Reservoir* in order to store (and disseminate) data by *making use of the (appropriately adapted/extended) EURAKNOS data model*.

¹² According to Wikipedia (), an Application Programming Interface (API) is “a computing interface to a software component or a system, that defines how other components or systems can use it” by defining the “kinds of calls or requests that can be made, how to make them, the data formats that should be used, the conventions to follow, etc.”

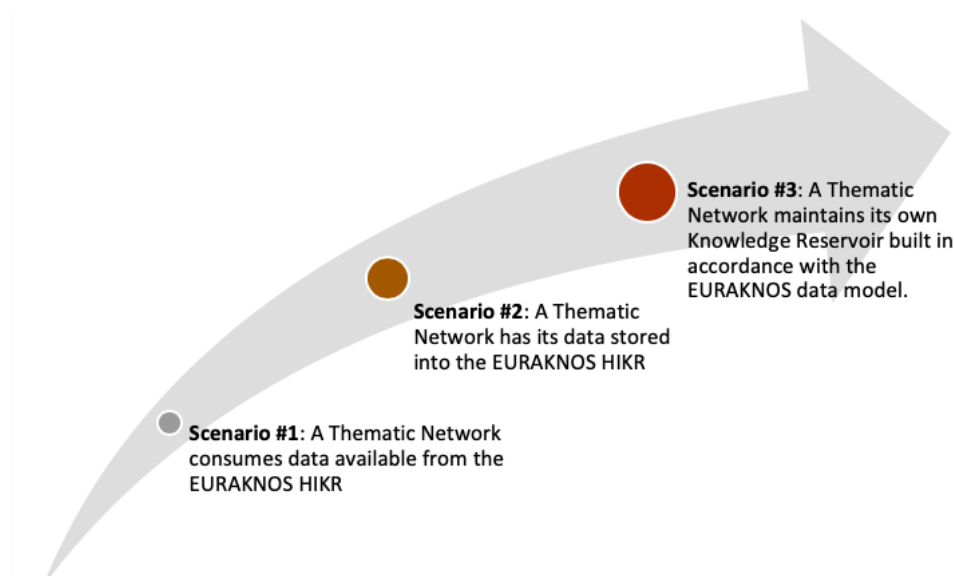


Figure 3: Scenarios illustrating the potential contribution of the EURAKNOS data model and repository to future Thematic Network efforts

4.2.3 Towards an ontology for formally describing digital outputs of Thematic Networks

Storing digital outputs of Thematic Networks efficiently in the EURAKNOS data repository (i.e. the HIKR) requires a robust data model illustrating output-related entities, entity attributes, as well as relations between them. These model elements will help determine an extensive list of metadata for a rich description of Thematic Network digital outputs. The EURAKNOS data model, which is core to the data repository's design, takes the form of an ontology presented in the following section. Despite the differences observed along the list of existing Thematic Networks with respect to a number of dimensions (domain, theme, aim and objectives, format and types of content, audience targeted), the EURAKNOS data model intends to provide a minimum set of requirements for efficiently describing the digital outputs available.

Within this context, the aim of this section is to outline the process towards the development of the EURAKNOS data model. The EURAKNOS data model is the outcome of the process illustrated in Figure 4 below.

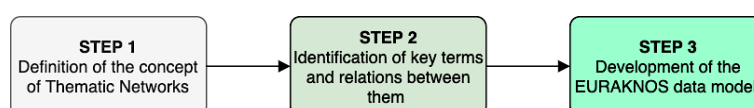


Figure 4: Methodology adopted for the work undertaken towards the development of the EURAKNOS data model

Definition of the concept of Thematic Networks

The basic assumption is that the effective identification of the data model's elements (namely, entities, entity attributes, and relations between entities) necessitates an explicit definition of the concept of Thematic Networks. Thematic Networks constitute a particular format of Multi-Actor projects, promoted by the agricultural European Innovation Partnership (EIP-AGRI) and funded by EU's Horizon 2020 programme, which focus on a specific theme. They bring actors from both science and practice together with the aim to collect and disseminate useful, practical outputs. Thematic Networks have stemmed from the need to efficiently address agriculture and forestry related problems and identify innovative solutions and best practices. However, solutions need to occur as the outcome of

coordinated, joint efforts of various actors bringing different perspectives and values to the process and, thus, contributing to synergies between research and practice. This approach is core to what EIP-AGRI calls an “*interactive innovation model*”¹³ and constitutes the foundation of Multi-Actor projects (EIP-AGRI, 2017).

Thematic Networks are a particular format of Multi-Actor projects (EIP-AGRI, 2016) and as such they:

- relate to specific themes (e.g. sustainable cropping systems, animal production systems, plant health, rural dynamics and policies, knowledge and innovation systems) and focus on real problems of practitioners;
- bring together partners with complementary backgrounds and expertise to collaborate, throughout the entire project lifecycle, and come up with innovative solutions;
- collect, further develop, as well as disseminate ready-to-apply recommendations and practices and, thus, serve as multipliers of Agriculture-related knowledge.

These distinctive features of Thematic Networks have provided the necessary input for developing the following, integrated definition of the concept of Thematic Networks:

Thematic networks are multi-actor projects that focus on a **theme**. The theme of a Thematic Network relates to a **domain**. Thematic Networks aim to disseminate easy to understand and apply information in the form of useful, practical **outputs** (practice abstracts, manuals, reports, infographics, videos, etc.). Outputs available by a Thematic Network can be distinguished into documents, software applications, presentations, images, audio and video objects, and datasets. Outputs are made available to end users (e.g. farmers, foresters, advisors) and have specific **types** and **formats**. Each output has a **purpose** (i.e. a reason for which it has been created), a **topic** (relevant to the Thematic Network’s domain) and has been created by a **creator**. Creators can be distinguished into **organisations** and **individuals** (who are affiliated with organisations). Organisations are involved in Thematic Networks either as **leading** or **participating** partners.

Identification of key terms and relations between them

A number of statements, helping to identify entities, entity attributes and relations between entities, can be derived from the above definition. In these statements, potential entities and entity attributes are highlighted in bold characters, whereas potential relations are highlighted in italicised characters.

- A **Thematic Network** *is about* a **theme**, and the **theme** *relates to* a **domain**. Thus, a **Thematic Network** *relates to* a **domain**.
- An **output** (available by a Thematic Network) is intended to be *used by* (specific **types of**) **end users**.
- There are different *kinds of* outputs: documents, software applications, presentations, audio and video objects, images, and datasets.
- An output *has* a {topic, purpose, type, format}.
- An output is *created by* a creator. A creator *is* an individual or organisation¹⁴.
- An **individual** is *affiliated with* an **organisation**.
- An organisation has a specific type of involvement in the Thematic Network (i.e. participating or leading partner).

It needs to be stressed that Thematic Network outputs are going to be stored and made available as digital files. Based on the work that has been undertaken in Task 2.1 (Mosquera Losada et al., 2019), and given the range of existing digital file types, the kinds of outputs that can be made available by Thematic Networks¹⁵ and are proposed in the EURAKNOS data model (i.e. documents, presentations,

¹³ <https://ec.europa.eu/eip/agriculture/en/eip-agri-concept>

¹⁴ This statement can also be formulated as follows: “An **individual** and an **organisation** are *kinds of* **creator**.”

¹⁵ Not only existing Thematic Networks, but future Thematic Networks as well.

images, software applications, audio and video objects, datasets) have been derived from the list of media types (termed as Multipurpose Internet Mail Extensions or MIME types) defined by the Internet Assigned Numbers Authority¹⁶. Media types in this list are: `application`, `audio`, `example`, `font`, `image`, `message`, `model`, `text` and `video`. Based on the research undertaken by Mosquera Losada et al. (2019), the media types that are considered relevant for determining the kinds of Thematic Network outputs are: `application`, `audio`, `image`, `text` and `video`. `Presentation` and `dataset` have been added to the media types subset considered. This analysis has resulted in a pool of terms, representing ontology-related entities or entity attributes, as well as a list of relations between terms. The terms and relations that have been derived from the definition of Thematic Networks are presented in Table 1 below.

Table 1: Potential ontology entities, entity properties and relations between entities

| Thematic Network – related terms (potential entities and entity properties) | relations between terms | | |
|--|-------------------------|--|--------------------------------|
| | relation label | relation domain ¹⁷ | relation range ¹⁸ |
| Domain ¹⁹ Topic Output Output type End-user Creator Format Purpose Individual Organisation Thematic Network | relates to | Thematic Network | Domain |
| | used by | Output | End-user (type) |
| | kind of | {Individual, Organisation} | Creator |
| | | {Document, Software application, Audio/Video object, Presentation, Image, Dataset} | Output |
| | Has | Output | {Topic, Type, Format, Purpose} |
| | created by | Output | Creator |
| | affiliated with | Individual | Organisation |

The EURAKNOS data model, provided in the form of an ontology, has been developed as the result of the analysis made on the definition of the Thematic Networks concept. Details on the model are made available in the following section.

4.2.4 Presentation of the EURAKNOS data model

The EURAKNOS data model takes the form of an ontology (Figure 5 below) that captures a number of entities of relevance to digital outputs of Thematic Networks, attributes of these entities, and the way in which these entities link to each other. The data model presented provides an explicit illustration of how data structuring has been conceived in EURAKNOS by taking account of the diversity of digital output types and formats available by Thematic Networks. It serves as the basis for the provision of a more elaborate data repository model, showing how data will be persisted into the EURAKNOS data repository, as well as for determining the metadata elements required for sufficiently and efficiently describing the digital outputs of Thematic Networks.

Ontologies make use of specific modelling conventions and elements indicative of the way in which these structures manage to represent knowledge. A core model element used by ontologies is that of

¹⁶ <https://www.iana.org/>

¹⁷ The term “**domain**” refers to the “source term” of the relation.

¹⁸ The term “**range**” refers to the “destination term” of the relation.

¹⁹ The term “**Domain**” refers to the wider Agriculture-related “area” that a Thematic Network relates to.

a **class**. Classes are abstract constructs used to denote collections of individuals/objects/concepts that share similar attributes. For instance, the class `Output`, which is used in the EURAKNOS data model, represents **any digital output** that could potentially be made available by a Thematic Network, rather than a specific data object. In the diagrammatic illustration of the EURAKNOS data model, in Figure 5 below, classes are represented as rectangles. The classes of the EURAKNOS ontology are: `Output` (with its subclasses `Audio`, `Dataset`, `Document`, `Image`, `Presentation`, `Software`, and `Video`), `Creator` (with its subclasses `Individual` and `Organisation`), `EndUserType`, `Topic`, and `ThematicNetwork`. An **instance** is a specific manifestation of a class; in other words, a concrete representative of that class. The **SMART-AKIS** Thematic Network is, for example, an instance of the `ThematicNetwork` class.

Class **properties** are also a key model element. We distinguish between two types of properties; namely, the **datatype properties** and the **object properties**. Datatype properties are properties that receive data values (e.g. strings, integers, double-precision floating point numbers). On the other hand, object properties are properties whose values are class instances. In the context of ontologies, object properties are conceptualised as **relationships**. In the EURAKNOS data model, relationships (i.e. object properties) are represented as edges connecting two rectangles. The use of black triangles is a design convention aiming to show the relationship's direction. For example, the `createdBy` relationship links the `Output` class with the `Creator` class with a direction from the `Output` class to the `Creator` class. So, based on the conventions used and the respective notations, the specific relationship may be read as: `Output` is `createdBy` `Creator`.

For reasons of simplicity and comprehensibility, the diagrammatic illustration of the EURAKNOS data model (Figure 5 below) contains only the object properties of ontology classes. Detailed descriptions of class properties (both datatype and object properties) are provided in Annex 1.

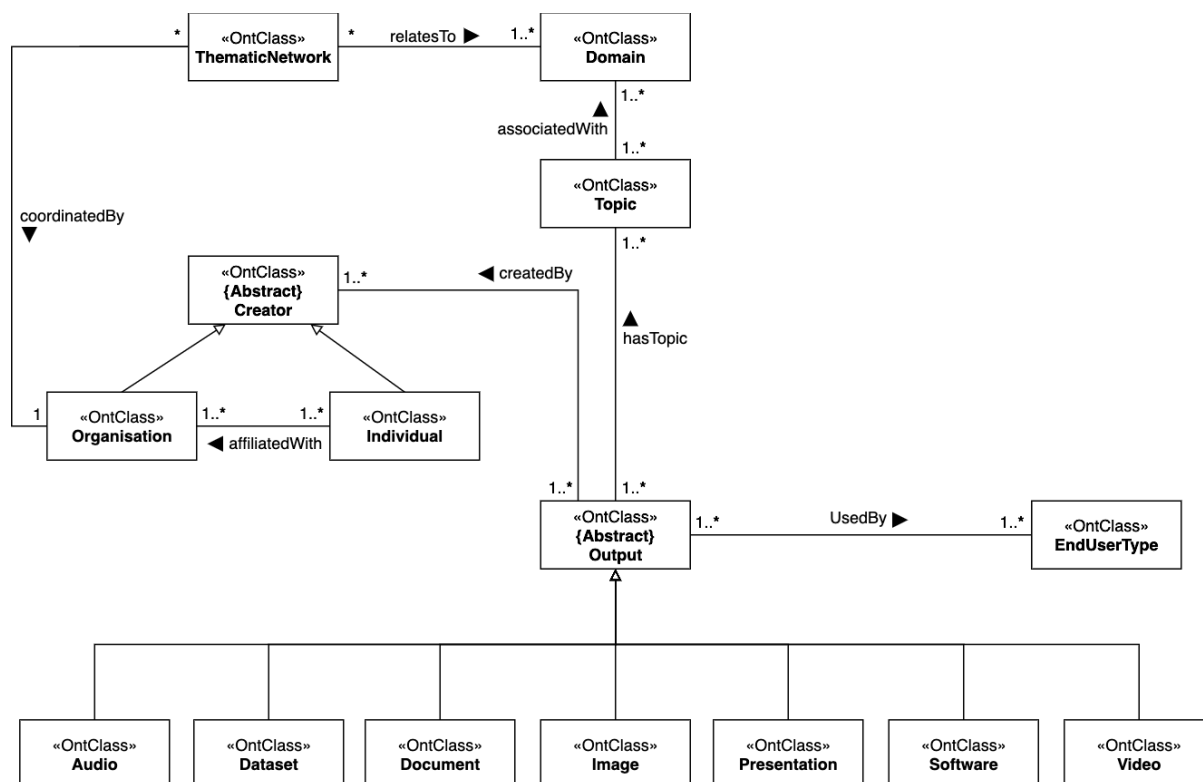


Figure 5: The EURAKNOS data model

An additional piece of information that is made available in the context of the EURAKNOS data model relates to **cardinality** or **multiplicity**. Cardinality/multiplicity is provided in the form of a range from a

minimum to a maximum value²⁰ and is used to inform about the number of class instances that may be associated with exactly one instance of the opposite side. So, one or more instances of the `Individual` class may be affiliated with one or more instances of the `Organisation` class and vice versa.

A specific type of relationship is the one that links a **subclass** to its **parent class**. Specifically, subclasses are linked to their parent classes by an “**is-a**” relationship. The notation used for this relationship type is an arrow with a non-filled triangle at its end linking the subclass to its parent class. So, for example, the `Document` class is a subclass of the `Output` class. The “**is-a**” relationship is associated with the concept of **inheritance** and, thus, has very powerful semantics. Inheritance means that all properties of the parent class are passed onto its subclasses. Based on the example above, all properties of the `Output` class are also properties of the `Document` subclass. Yet, the `Document` subclass has some additional properties.

A point that raises attention is that some classes (namely, the `Output` and `Creator` classes) have been labelled as **Abstract**. Abstract classes do not have direct instances; only their subclasses do. So, a specific individual that has created a specific output is an instance of the `Individual` subclass and is not considered to be a direct instance of the `Creator` class.

The diagrammatic representation of the EURAKNOS data model has been based on the use of model elements of a **UML (Unified Modelling Language)**²¹ **class diagram**. Class diagrams are tools commonly employed for the needs of illustrating ontologies, because of a number of similarities in the modelling conventions employed. However, there are significant differences as well. Therefore, **Ontology UML Profiles** are extensions of the set of design elements and conventions used in a UML class diagram, which have been proposed²² to efficiently address ontology-related particularities. The EURAKNOS data model has been created by making use of a bare minimum set of UML-related model elements so as to enhance the interpretability and comprehensibility of the design. For the same reason, no particular UML profile design add-ons have been employed except from the «`OntClass`» tag that is used to denote that the model’s classes are ontology classes.

4.2.5 Compliance with the FAIR principles

An issue that is important to consider in the Big Data era is that every publisher can publish information through a data format that is only processable by specific systems. Thus, the resulting data ecosystem becomes less integrated from day to day and this fact that has significant implications with regard to the discovery and reuse of valuable resources. In order to achieve sustainable data management, the data community has created a set of high-level, technology-agnostic principles known as the **FAIR Data Principles** (Wilkinson et al., 2016). These Principles describe distinct considerations for data publishing to make digital resources **Findable, Accessible, Interoperable, and Reusable**. Moreover, because of the information overload experienced by human agents, these principles also emphasise the capacity of computational systems to find, access and reuse data with minimal human intervention. EURAKNOS will undertake these core principles by: (i) using standard metadata schemas to enable findability; (ii) using open standards such as HTTP and RDF to maximise accessibility; (iii) linking data with other FAIR datasets to improve interoperability; and (iv) making use of rich metadata sets with the aim to enhance reusability.

²⁰ It is worth to be mentioned that the * cardinality-related notation is equivalent to the **0..*** notation (which means greater than or equal to zero).

²¹ **UML (Unified Modelling Language)** is a general-purpose, developmental, modelling language used in the field of software engineering, which is intended to provide a standard way to visualise the design of a system (Booch et al., 2005).

²² See, Baclawski et al. (2001) and Djurić et al. (2004) for reference.

4.3 Considerations for selecting data objects to store into the High Impact Knowledge Reservoir

The data repository (i.e. High Impact Knowledge Repository) of the prototype electronic Knowledge Reservoir Platform (e-KRP) under development in EURAKNOS will store data objects from a number of eight (8) Thematic Networks. To this end, before proceeding to a selection of Thematic Networks from which data objects will be collected and integrated into the HIKR (an activity to be undertaken as part of Task 4.2), there is need to identify the **types of data objects** and **formats of content available** from them that best fit the needs of the end-user types targeted. Farmers, foresters and advisors have been identified as the core end user types of the e-KRP. However, in the context of the analysis of platform-related initiatives, available at both the international and national levels, undertaken by Gernert et al. (2019), educators and policy makers have also been identified as potential user categories. Conclusions about the types of data, and format of content provided, to be considered and integrated into the HIKR of the prototype platform have been based on the work of Tasks 3.1 and 4.1. These tasks have been concerned with systematic efforts to produce and provide alignments between the needs of end-users and the type/format of data objects and content that can help to best meet these needs. More specifically, Task 3.1 deals with the best methodologies and practices to produce, select, collect and store data as outputs of Thematic Networks and is also concerned with the identification of the types of data objects (and formats of content provided by them) that should be integrated into the EURAKNOS HIKR. Task 4.1 has been concerned with the investigation of the feasibility and added value of a new, centralised agriculture-related knowledge platform. For this purpose, a systematic mapping and analysis of existing platform-related initiatives has taken place with a focus on knowledge-related needs of end users. The process from the identification of the e-KRP end-user types to the integration of data objects from (eight) Thematic Networks into the HIKR is shown with the help of the workflow presented in Figure 6: Process from the identification of the e-KRP end-user types to the integration of TN data into the EURAKNOS HIKR below.

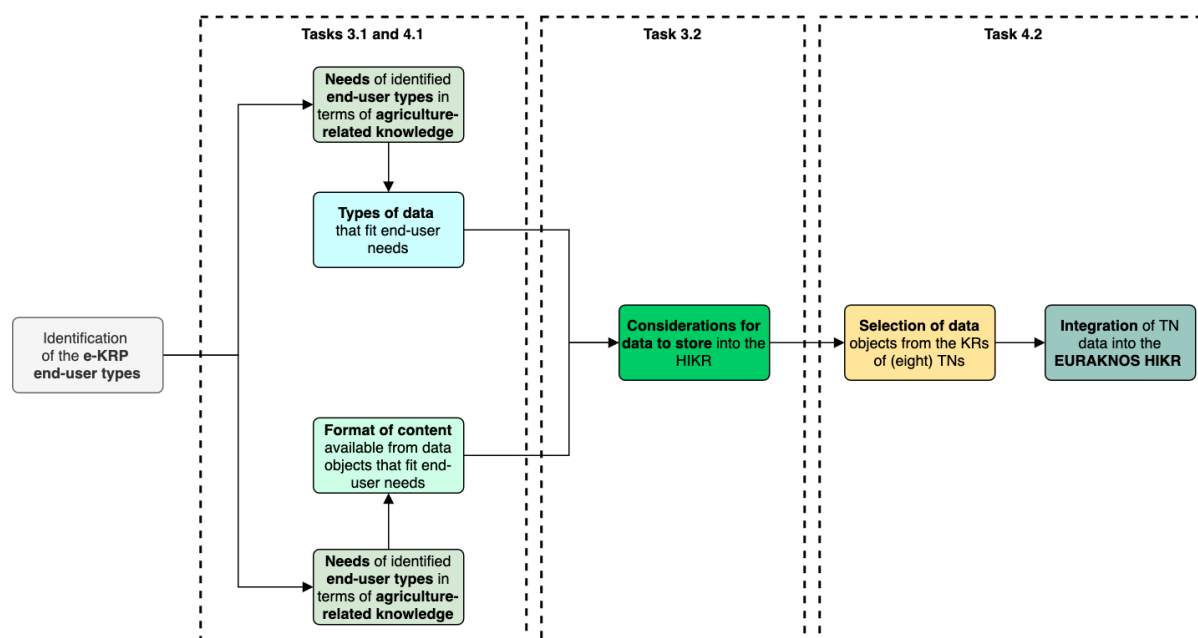


Figure 6: Process from the identification of the e-KRP end-user types to the integration of TN data into the EURAKNOS HIKR

The EURAKNOS meeting in Paris (12th to 13th December, 2019) hosted a workshop dedicated to the aim and activities of Task 3.1 (Working Group 3.1 workshop) in the context of which, the experts invited provided their inputs regarding the evaluation of the data available by the Knowledge Reservoirs of

Thematic Networks and contributed to the development of a list of criteria for the content to be made available through a HIKR (i.e. a Knowledge Reservoir that provides content of importance to the users targeted).

According to the workshop outcomes reported by Marois et al. (2020), a HIKR should primarily target at the needs of farmers and advisors who have been identified as major end-user types. Based on that, there needs to be a core focus on the provision of data objects conveying practical and ready-to-apply knowledge. However, the information available to end-users should also reflect research findings that have been appropriately tested and validated. **Booklets, videos and brochures** have been proposed as kinds of data objects that should be opted for the delivery of content and information to the end-users with videos being considered as more appealing to farmers. Content produced by peers is highly valued by farmers and, therefore, should be provided by a HIKR. Making content and information available in languages other than English has also been brought up as an issue of importance.

The need to provide farmers with ready-to-apply, practical and comprehensible information has also been highlighted by Gernert et al. (2019), who have proceeded to an alignment between different types of end-users and their needs in terms of agriculture related knowledge (shown in Table 2 below).

Table 2: Alignment between end-user types and their needs in terms of Agriculture-related knowledge (adapted from Gernert et al., 2019)

| | characteristics of knowledge needed | Characteristics of content needed | types/formats of data conveying the knowledge and content required |
|--------------|--|--|---|
| Farmer | <ul style="list-style-type: none"> - easily accessible - comprehensible - useful, practical - up-to-date | <ul style="list-style-type: none"> - tailor-made advice - technical knowledge to be applied in the farm - able to be adapted to their context and needs | <ul style="list-style-type: none"> - practice abstracts - video-based material - factsheets - newsletters - training material - farmer magazines/journals |
| Advisor | <ul style="list-style-type: none"> - tailor-made - comprehensive - transferable - up-to-date | <ul style="list-style-type: none"> - recent policy- and sector-related developments - localised solutions for specific business and technical problems | <ul style="list-style-type: none"> - practice abstracts - newsletters - factsheets - press releases - video-based material |
| Educator | <ul style="list-style-type: none"> - up-to-date - transferable | <ul style="list-style-type: none"> - scientific results and practice - adapted to the target group and practitioner needs | <ul style="list-style-type: none"> - practice abstracts - factsheets - training courses/course modules - teacher guides - resource packs |
| Policy maker | <ul style="list-style-type: none"> - standardised/generalised - evidence-based - up-to-date - applicable | <ul style="list-style-type: none"> - relevant to their respective policy area - design of new policies | <ul style="list-style-type: none"> - press releases - video-based material - newsletters - factsheets - policy reports & recommendations |

As made evident from the table above, **practice abstracts, factsheets, newsletters and press releases**, as well as **video-based material** constitute types and formats of data that have been identified as the

most relevant to the needs of farmers and advisors. In their majority, these types and formats are also appropriate for addressing the needs of educators and policy makers.

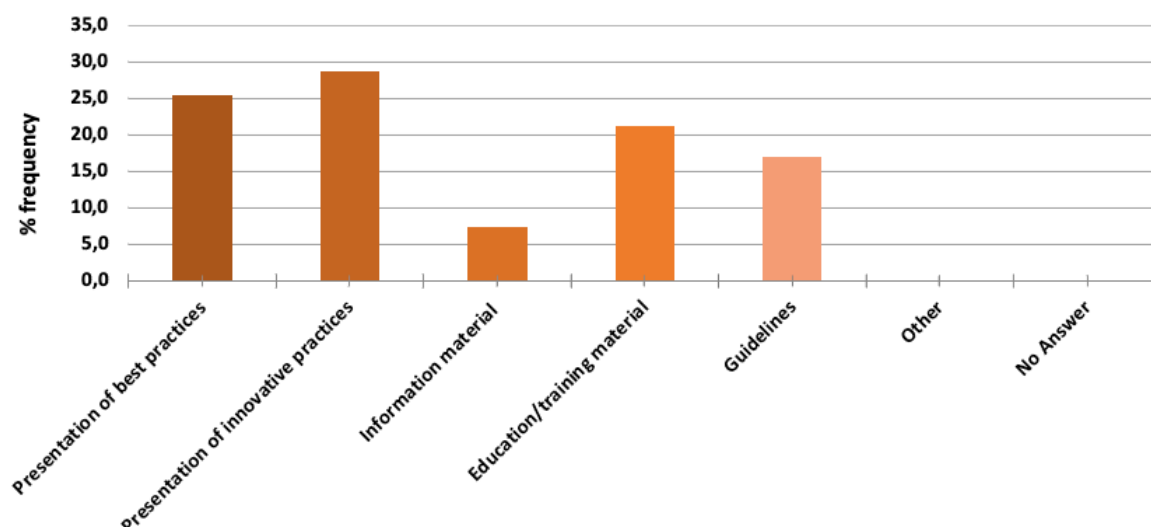


Figure 7: Advisor responses to the question “What kind of content/information would you prefer?”

In order to strengthen the insights provided by Tasks 3.1 and 4.1, decisions on the types of data and content formats, to be considered for the EURAKNOS HIKR, have also been based on the results of the advisor survey²³ that took place in the context of the EUFRAS meeting²⁴ in Athens. The survey took place on the 25th of February 2020, and was implemented by administering questionnaires to a sample of 32 advisors. From the survey questions, those of relevance to Task 3.2 were the following:

- **Q1:** “What kind of content/information would you prefer?”
- **Q2:** “Which of the following formats of digital content is most preferable to you?”

As shown in Figure 7 above, presentation of innovative and best practices are the most popular kinds of information/content among advisors by receiving a response rate of 28.7% and 25.5% respectively. Training and educational material is also important, by receiving a response rate of 21.3%, followed by content related to guidelines with a response rate of 17%. Information material has been rated as the least preferred kind of content/information (response rate = 7.4%). For reasons of clarification, it needs to be mentioned that the respondents were able to select more than one options related to the kinds of information and content they prefer.

With regard to Q2, video has been rated as the most preferable format of digital content, by receiving a response rate of 27.1%. Video is followed by presentation- (25.7%) and text-based content (22.9%). Images (14.3%) and audio-based material (7.1%) are reported as the least preferable formats of digital content for advisors. Results that have been obtained with regard to Q2 are shown in Figure 8 below.

²³ More details with regard to the advisor survey are provided in Annex 2.

²⁴ <http://www.eufiras.eu/index.php/activities/archive/142-eufiras-annual-assembly-2020>

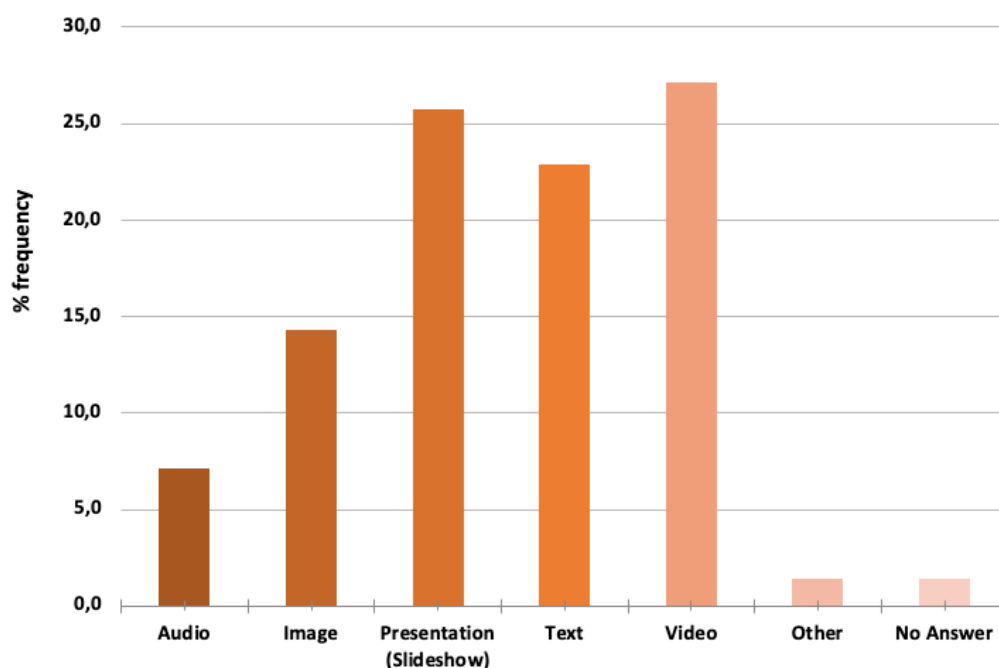


Figure 8: Advisor responses to the question “Which of the following formats of digital content is most preferable to you?”

The EURAKNOS data model distinguishes among seven **kinds of digital outputs** of Thematic Networks (namely, “audio”, “dataset”, “document”, “image”, “presentation”, “software”, and “video”)²⁵, each of which is associated with a list of **formats** (i.e. file formats of digital outputs per kind) and **types**. Type is tightly associated with the content and information conveyed by a data object and can be considered indicative of the **purpose** that the data object (i.e. the output of the Thematic Network) serves. Types per kind of output have been identified by drawing on the work of Mosquera Losada et al. (2019) and by trying to have a view on the content that may potentially be provided by future Thematic Networks. In order to better illustrate the preferences of advisors with regard to Agriculture-related content and information, Table 3 below provides an alignment between advisors’ preferences²⁶ and the types of digital outputs of Thematic Networks that, according to the EURAKNOS data model, are most likely to match the preferences reported.

Table 3: Alignment between advisors’ preferences and the most relevant kinds and types of outputs proposed in the data model

| preferred kind of content/information | type of output (based on the EURAKNOS data model) | kind of output (based on the EURAKNOS data model) |
|--|--|---|
| presentation of innovative/best practice | Guide | audio, document, presentation, video |
| | article in conference proceedings, book, chapter in edited volume, deliverable report, journal article, report/paper | document |

²⁵ These kinds of digital outputs are modelled with the help of the Audio, Dataset, Document, Image, Presentation, Software, and Video classes.

²⁶ In order to focus on the most preferable kinds of content/information, those rated below 10% (i.e. “**information material**”) have not been considered.

| | | |
|-----------------------------|---|---|
| education/training material | educational/training podcast, on-demand seminar handbook educational/training presentation webinar, educational/training video | audio document presentation video |
| guidelines | guide tutorial manual, technical article | audio, document, presentation, video document, presentation, video document |

Considerations and insights with regard to the types of data, and formats of content conveyed, that could be of relevance to the end users targeted, can be summed up to the following points:

- Content and information delivered through video-based data objects is preferred the most by all types of end users targeted.
- There is a range of content and information provided in text-based formats (practice abstracts, factsheets, booklets, handbooks, brochures, guides, tutorials, manuals, education and training material) that has the capacity to cover the needs of the end users targeted; documents are a kind of digital output that is highly valued by both farmers and advisors.
- Evidence for preferences in slideshow-based content/information (i.e. presentations) has been made available only for advisors. Advisors are more in favour of slideshow-based material than content and information delivered in textual formats.
- Preferences in image- and audio-based content and information have been reported only for advisors. By taking account of their preferences in content and information related to best and innovative practices, as well as guidelines, it can be concluded that audio-based material (e.g. podcasts) conveying such kinds of information would be relevant to advisors' needs.

On top of the conclusions listed above, there are some additional points to consider:

- There is a requirement to have data from eight (8) Thematic Networks stored in the repository of the prototype platform as a proof of concept for the design and development made.
- Some combinations between kinds and types of digital outputs that are preferred by the user types targeted may not be available.
- Evidence with regard to preferences in slideshow-based (i.e. presentations), image-based, and audio-based content/information is only available for advisors. However, we may assume that these kinds of digital outputs could also be of relevance to farmers (and foresters).
- The vast majority of digital outputs provided by Thematic Networks relates to types associated with text-based formats (e.g. practice abstracts, factsheets, press releases, research/technical articles, handbooks, reports, and guides). Video-based content and information is available at a lower rate (it counts for 41.7% of the digital outputs available). The same applies to image-based content/information and podcasts (i.e. audio-based content/information), which count for 33.3% and 20.8% of the total content respectively (Mosquera Losada et al., 2019).

Within this context, the sample of digital outputs to be integrated into the repository of the prototype e-KRP will contain:

- 1) **videos** of any type available;
- 2) **documents** with an emphasis on types like practice abstracts, factsheets, tutorials, manuals, guides, handbooks, technical articles, booklets, brochures, newsletters, press releases, journal articles, chapters in edited volumes, articles in conference proceedings, deliverable reports, reports/papers, and books²⁷;

²⁷ By taking account of the fact that articles in conference proceedings, books, chapters in edited volumes, deliverable reports, journal articles, and reports/papers are not the most preferable types of data by the e-KRP end-user types, only a few data objects from this subset will be collected and stored into the HIKR.

- 3) **presentations**: there will be an attempt to be emphasise upon types like guides, tutorials, and educational/training presentations. In the case that such types of presentation-based outputs are not available, then any type available will be collected and integrated in the HIKR.
- 4) **images** of any type available;
- 5) **audio-based outputs**: there will be an attempt to emphasise upon types like educational and training podcasts, on-demand seminars, and guides. In the case that such types of audio-based outputs are not available, then any type available will be collected and integrated in the HIKR.

The ranking of the output kinds presented above is indicative of priorities and the number of items per output kind, to be stored in the HIKR, in relation to the other kinds considered (see Figure 9 below).

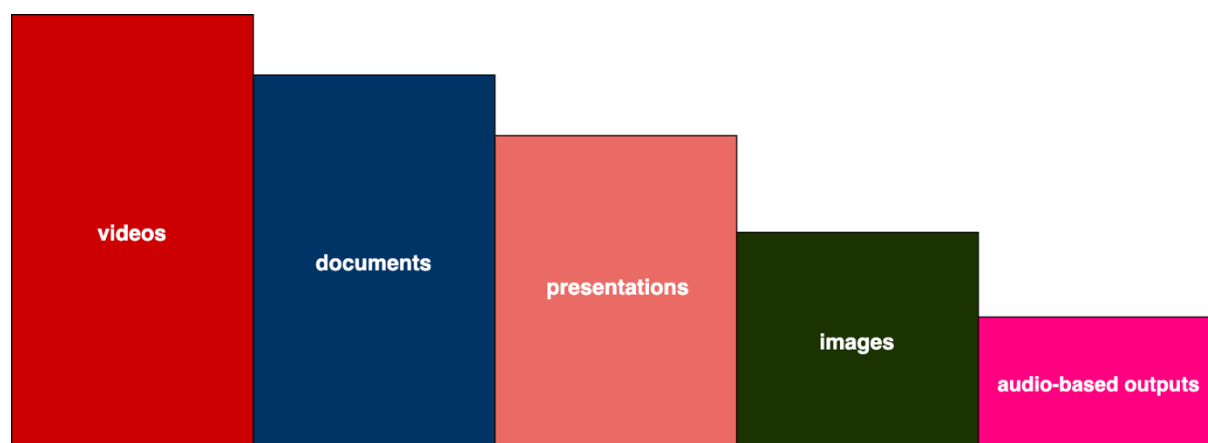


Figure 9: Kinds of Thematic Network outputs to be integrated into the repository of the prototype e-KRP; the box heights are indicative of number of items per output kind to be considered in relation to the other kinds

4.4 Technology framework

The solution that has been considered and adopted for the needs of the development of the HIKR and, thus, the e-KRP of EURAKNOS is **DSpace**, an “out-of-the-box open source software package for creating repositories focusing on delivering digital content to end users and providing a full set of tools for managing and preserving content within the application”²⁸. DSpace does not just provide digital file preservation and storage functionalities, but rather goes beyond that and provides functional preservation according to which files are “kept accessible as technology formats, media and paradigms evolve over time for as many types of files as possible”²⁹. DSpace is a full-stack web application comprising a database, storage manager, and a front-end web interface.

The reason for selecting DSpace is that it constitutes the most widely adopted data repository creation solution with more than 2,000 installations worldwide³⁰. It has a vibrant community built around it, it is flexible, and it has the capacity to provide solutions to academic and research institutions as “an open access institutional repository for managing and disseminating scholarly output”, as well as to other types of organisations for the needs of “hosting and managing subject-based, dataset or media-

²⁸ <https://duraspace.org/dspace/resources/technical-specifications/>

²⁹ https://duraspace.org/wp-content/uploads/dspace-files/DSpace_Diagram.pdf

³⁰ <https://duraspace.org/dspace/resources/technical-specifications/>

*based repositories*³¹. It needs to be noticed that integration with OpenAIRE is one of DSpace's core features³². A number of project cases that make use of DSpace³³ are listed below:

- **Deep Blue**³⁴ is the University of Michigan's service for providing access to content related to research, teaching, and creativity, which makes use of DSpace as a repository for a range of digital objects including journals.
- Rice University's **TIMEA** initiative³⁵ is a digital archive regarding Western interactions with the Middle East, particularly travels to Egypt during the nineteenth and early twentieth centuries. By making use of a repository built with the help of DSpace, TIMEA allows access to digitised content such as travel guides and narratives, museum catalogues, historical maps, interactive GIS maps of Cyprus and Egypt, as well as photographic and hand-drawn images of Egypt.
- **Policy Archive**³⁶ is a digital library of public policy research containing over 24,000 documents. By making use of a repository, developed and maintained with DSpace, the institution provides a digital archive of global, non-partisan public policy research that makes use of the Internet technology to collect and disseminate content related to policy research available in the form of text, video, reports, briefs, and multimedia material.

Other notable cases of open data repositories built on DSpace are the Open Knowledge Repository of World Bank³⁷, the Apollo repository of the University of Cambridge³⁸, the Digital Access to Scholarship initiative of Harvard³⁹, DSpace@MIT⁴⁰, the Spiral repository of the London Imperial College⁴¹, and the Institutional Repository for Information Sharing of the World Health Organisation (WHO)⁴².

DSpace is the "umbrella" technology solution that is intended to be employed in EURAKNOS. However, there are specific issues that need to be addressed for the development of the e-KRP architecture. Such issues relate to data and metadata storage, as well as search engine functionalities. Given the fact that a detailed presentation of the architecture of the EURAKNOS e-KRP (in-progress) is beyond the scope of this deliverable⁴³, some initial considerations regarding data/metadata storage and potential search engine system solutions are provided in the following sub-sections.

4.4.1 Data storage

A core component of the DSpace framework is PostgreSQL, an object-relational database used for the storage of items (and metadata). According to the official documentation⁴⁴, PostgreSQL is a highly scalable, SQL-compliant, open source, object-relational database management system with some of its features being the following:

- **Open-source community:** PostgreSQL is the result of a process of fifteen years of development with version 12.2 being the latest one.
- **Advanced SQL:** PostgreSQL supports complex queries and ANSI SQL 2003.
- **Extensibility:** PostgreSQL supports a number of advanced data types such as full text and JSON.

³¹ <https://duraspace.org/dspace/resources/technical-specifications/>

³² <https://wiki.lyrasis.org/display/DSPACECRIS/OpenAIRE+Integration>

³³ More information about DSpace use cases is available at: <https://duraspace.org/dspace/about/use-cases/>

³⁴ <https://deepblue.lib.umich.edu/>

³⁵ <http://timea.rice.edu/>

³⁶ <https://www.policyarchive.org/>

³⁷ <https://openknowledge.worldbank.org/>

³⁸ <https://www.repository.cam.ac.uk/>

³⁹ <https://dash.harvard.edu/>

⁴⁰ <https://dspace.mit.edu/>

⁴¹ <https://spiral.imperial.ac.uk/index.jsp>

⁴² <https://apps.who.int/iris/>

⁴³ Details on the architecture of the EURAKNOS e-KRP will be provided in D4.3 ("EURAKNOS e-KRP documentation and guide").

⁴⁴ https://wiki.postgresql.org/images/4/43/What_is_pg%28%29.pdf

- **Freedom:** PostgreSQL comes with a BSD-licensed that allows to use it for any kind of purpose including commercial exploitation.
- **Languages:** PostgreSQL has an interface with stored procedures developed in widely adopted programming languages.
- **Concurrency:** Multiple users can concurrently work with the database.

4.4.2 Metadata storage

OpenLink Virtuoso⁴⁵ is a multi-model database that supports, among others, the management of data represented as relational and graph models and is considered as solution for metadata storage in the context of EURAKNOS. It supports different RDF serializations such as N3 and RDF/XML, as well as the SPARQL Query Language. By being a graph database, OpenLink Virtuoso stores data as a network of objects connected with links between them. This makes OpenLink Virtuoso (and RDF triplestores in general) a choice to consider for managing highly interconnected data, such as metadata. Linked Open Data datasets, like DBPedia and GeoNames, are published, for example, as RDF triplestore databases. This allows for obtaining highly relevant search results from federated queries much faster. Moreover, this type of databases is also capable of handling complex semantic queries and of using inference for uncovering implicit information out of the existing (explicit) relations.

4.4.3 Search engine

There is a number of options that have been examined with regard to the search functionality of the EURAKNOS HIKR. These options have not been considered with the aim to replace the search functionality built in DSpace, but rather to complement it. Relational Database Management Systems (RDBMS), such as MySQL, SQLite, PostgreSQL, Microsoft SQL Server, MariaDB etc., are the cornerstone of data persistence in software development. Their longevity, portability, well-written Graphical User Interface tools, and ease of querying, make them the most popular choice for the deployment of data storage mechanisms. The representation of data in a tabular form (i.e. rows of records and columns of fields) is an intuitive way to organise many transactional data types. As a consequence, it is natural to consider the relational model when trying to organise data for a search application. However, modern applications characterized by increased data workloads may suggest that RDBMSs are not scoring very high with regard to search functionality, due to scalability- and speed-related constraints.

Search can be enhanced by making use of **Solr**⁴⁶, which is the search technology shipped with DSpace. Solr is an open-source, enterprise-search platform developed in Java as part of the Apache Lucene project⁴⁷. Solr has the capacity to support the efficient querying of data and to provide adequate information to the calling component of the system (front-end application). The data stored in the underlying Lucene index will facilitate the search service aspect of the application by storing only searchable view of the data, residing in as a decoupled component of system. Some of Solr's major features are full-text search, hit highlighting, faceted search⁴⁸, real-time indexing, dynamic clustering, database integration, NoSQL features and rich document handling. By allowing for distributed search and index replication, Solr is able to provide scalability and fault tolerance. Solr is widely adopted and used for enterprise search and analytics use cases and has is supported by an active development community releasing new product version on a regular basis.

Solr is able to achieve fast search responses by searching an index instead of searching the text directly, which is analogous to making use of a keyword index in the end of a book. This type of index is called **inverted index**⁴⁹, because it inverts the page-centric data structure (page to words) to a keyword-centric data structure (word to pages). Regarding the representation of data, it is a document-based

⁴⁵ <https://virtuoso.openlinksw.com/>

⁴⁶ <https://lucene.apache.org/solr/>

⁴⁷ <https://lucene.apache.org/>

⁴⁸ <https://www.sciencedirect.com/topics/computer-science/faceted-search>

⁴⁹ https://en.wikipedia.org/wiki/Inverted_index

engine, meaning that a document is the basic unit of search and index. An index consists of one or more “**Documents**”, and a “**Document**” consists of one or more “**Fields**”. In database terminology, a “**Document**” corresponds to a table row, and a “**Field**” corresponds to a table column. Before adding documents to Solr, a schema needs to be specified and declare: (i) the kinds of fields available; (ii) the field to be used as the unique/primary key; (iii) the fields required; and (iv) how to index and search each field, as well as the specific field types (e.g. float, long double date text).

In order to search a document, Apache Solr performs the following operations in sequence:

- **Indexing:** Apache Solr converts the documents into a machine-readable format which is called “Indexing”.
- **Querying:** Relates to the understanding of the terms of a query posed by the user. These terms can be images or keywords.
- **Mapping:** Apache Solr maps the user query to the documents stored in the database in order to find and deliver the most appropriate result(s).
- **Ranking the outcome:** As soon as the engine searches the documents indexed, it proceeds to a ranking of the outputs in terms of relevance.

Apache Solr runs as a standalone full-text search server. It uses the Lucene Java search library at its core for full-text indexing and search, and has REST-like, HTTP/XML and JSON APIs that make it usable from the most popular programming languages. Documents can be inserted in it (“Indexing”) via JSON, XML, CSV or a binary format over HTTP, and queried via HTTP GET. Results can be delivered in JSON, XML, CSV or binary formats. A number of features that make Solr popular and are relevant for the EURAKNOS platform are the following:

- indexing in near real time;
- automated index replication;
- server statistics logging;
- automated failover and recovery;
- rich document parsing and indexing;
- multiple search indexes;
- user-extensible caching;
- design for high-volume traffic;
- scalability, flexibility and extensibility;
- advanced full-text searching;
- geospatial searching; and
- load-balanced querying.

Apache Solr’s basic architecture is shown in Figure 10 below and its major components are⁵⁰:

- **Request Handler** – The requests sent to Apache Solr are processed by these request handlers. The requests might be query requests or index update requests. Based on the requirements, the request handler must be selected. To pass a request to Solr, the handler is mapped to a certain URI end-point and the specified request is served by it.
- **Search Component** – A search component is a type (feature) of search provided in Apache Solr. It might be spell checking, query, faceting, hit highlighting, etc. These search components are registered as **search handlers**.
- **Query Parser** – The Apache Solr query parser parses queries that are passed to Solr and checks them for syntactical errors. Upon completion of parsing, it translates them to a format that Lucene can “understand”.
- **Response Writer** – A response writer is the component that generates the formatted output for user queries. Solr supports response formats such as XML, JSON, CSV, etc. Different response writers are available for each type of response.

⁵⁰ https://www.tutorialspoint.com/apache_solr/apache_solr_architecture.htm

- **Analyser/tokenizer** – Lucene recognises data as tokens. Apache Solr analyses the content, splits it into tokens, and passes these tokens to Lucene. An analyser examines the text of fields and generates a token stream. A tokenizer breaks the token stream, prepared by the analyser, into tokens.
- **Update Request Processor** – Whenever an update request is sent to Apache Solr, the request is executed through a set of plugins (signature, logging, indexing) known as **update request processor**. This processor is responsible for modifications such as dropping a field, adding a field, etc.

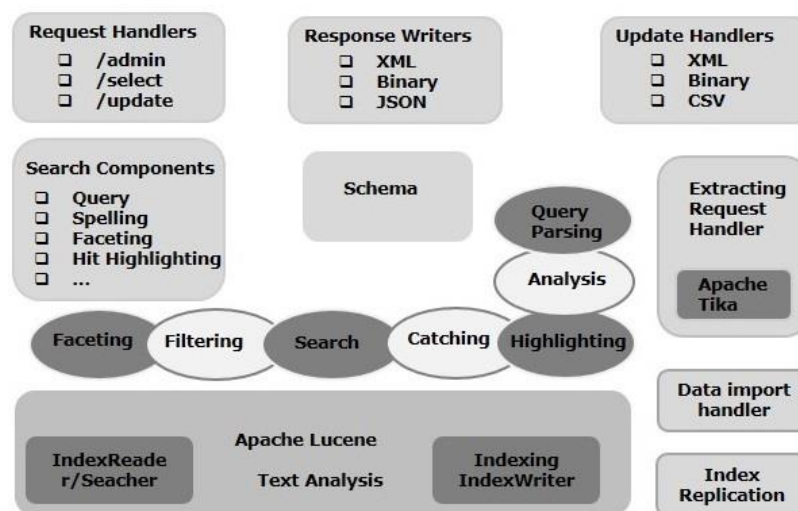


Figure 10: Basic architecture of the Solr search engine

Another search engine alternative considered, similar to Solr, is **Elasticsearch**⁵¹, which is also built on top of the Apache Lucene library⁵². Elasticsearch provides a distributed, multitenant-capable, full-text search engine. It is also developed in Java, it has an HTTP web interface, and it makes use of schema-free JSON documents. Since both search technologies are based on similar principles, providing good scalability and performance, there is no question of selecting the best one of two. Final decisions are going to be made in the development phase when the HIKR and prototype e-KRP are going to be built.

⁵¹ <https://www.elastic.co/elasticsearch/>

⁵² <https://www.oreilly.com/library/view/elasticsearch-the-definitive/9781449358532/>

5. From HIKR to e-KRP (Digital Platform)

5.1 Types of knowledge reservoirs in current TNs

If we recap the classifications of Knowledge Reservoirs (KR) in current Thematic Networks (see D2.2), we identified 4 main types + 1 'hybrid' category:

1. KRs as a catalogue, choose your innovation
2. KRs as a library: find the knowledge on the right shelves
3. KRs as a practical digital guide book
4. KRs as a market place to exchange good practices
5. KRs as a mix of different categories

The main functionalities of these types of KRs are:

1. KRs as a catalogue, choose your innovation
 - This type of KR supports the user in his choices oriented in the search for practical solutions. So, these KRs reference innovative technical solutions. The key steps for consulting content are in the first place through the use of a search engine; in the second place search criteria are defined to filter solutions.
2. KRs as a library: find the knowledge on the right shelves
 - The main functionalities identified are a classification of contents by theme or type of content. The key steps for consulting content area guidance on the site according to a proposed thematic classification
3. KRs as a practical digital guide book
 - The main functionalities identified are a presentation of tricks and tips and very practical solutions. The key steps for consulting content are first, through the use of a search engine, second search criteria defined to filter solutions or defined on the subject on which it is focused.
4. KRs as a market place to exchange good practices
 - The main functionalities identified are a KR which does not host any knowledge content but which lists references accessible through URL links.

5.2 A (preliminary) future vision for the e-KRP in HIKR

When we take these different types of Knowledge Reservoirs into consideration and the user requirements (see D2.5), we foresee an evolutionary path for a future digital platform (e-KRP).

Based on the recommendations in D2.2 & The KIP meeting in Budapest (D2.5), The e-KRP should cater (at least) to these objectives:

- A central database that can store data from multiple Thematic Networks (main aim of the
- A user interface that enables searching this data in an intuitive way
- A personalized 'user journey' for the visitor, in which content is tailored to the actual visitor.
- Include functionalities to gather insights on user needs & content feedback
- Include functionalities that enable improvements over time (eg. statistics)
- Availability of different languages

In the context of Euraknos we are building a prototype of the e-KRP. This prototype should pave the way for the full-fledged platform 'FarmBook' (workname) in the future. To enable the continuous

development of the future vision for the e-KRP, we suggest to focus first on the key functionality - a central database - and expand the scope of the platform in the coming years.

5.2.1 Phase 1: Prototype - LIBRARY

In phase 1 of the e-KRP: data from different thematic networks is stored in a central database and is easily searchable.

The main user journey is focused on easily searching relevant content in the database. The aim is a highly structured – functional – format for the platform with a clean and attractive appearance.

To cater to the philosophy of open access & the FAIR principles, the database will not require the user to create an account and login to the platform to search the database. All data is thus findable & accessible for the user.

The main disadvantage and direct consequence of an 'unregistered user' is that the level of personalization is limited. The visitor has to accept cookies in order to be able to allow any personalisation to be saved (2/3 of the cookies are automatically rejected by most web browsers). If the visitor deletes the cookies, the preferences and personalisation will be deleted as well. The visitor will however still be able to get suggestions for interesting content based on related tags (by contributors) and most viewed content within specific categories.

When we add a login module, the e-KRP allows visitors to register an account. The main advantage for the user in this state is that it allows the e-KRP to remember the user's preferences across sessions, and thus provide a more (automatically) tailored user journey.

The prototype will enable:

1. A central database for the data objects of at least 8 Thematic Networks
2. Direct access to content in the database (no login required)
3. Include a dashboard with a clean, attractive look
4. Searchable via filters (which can be combined)
 - New/old content
 - Most viewed
 - Theme (e.g.. precision farming, forestry,...)
 - Type of format (e.g.. video, audio, pdf,...)
 - Type of content (e.g.. practice abstract, fact sheet,...)
 - ...
5. 'Consume' content on the platform
 - Read articles, watch video, listen to audio,...
6. Download content from the platform
 - Download pdfs,...

5.2.2 Phase 2: Minimum Viable Product (MVP) – Platform

In phase 2 of the e-KRP (out of scope for the EURAKNOS project), the platform will build upon the central database and add new functionalities that cater to an expanded set of user needs.

The platform will enable:

1. Personalization after login
 - (eg. create an account, communicate & save preferences,...)
2. To gather basic insights on user needs & feedback on content
 - (eg. Usage statistics, general rating of content, express whether content was useful/relevant for me or not,...)
3. Recommendations for the end-user
 - (eg. Recommended content based on your search queries / preferences, people also looked for,...)
4. Alerts on relevant content
 - (eg. End users can set specific alerts for new content on a topic,...)
5. Understanding the ‘source’ of the content
 - (eg. Include the name of the contributing organization or ‘author’ of content,...)
6. Searching for contacts
 - (eg. Search organizations in my region, get contact details from authors or contributors,...)
7. ‘Basic’ sharing options of content to social platforms
 - (eg. Share to Facebook, LinkedIn,...)
8. Include features for accessibility
 - (eg. Increased contrast and font legibility,...)

5.2.3 Phase 3: Farmbook

In the third evolution of the platform, the database will still be the foundation, but the purpose of Farmbook will broaden to workspace & social platform.

In this phase the platform caters to user needs, such as:

1. Features that enable an interactive onboarding flow for new users on how to use the platform
 - (eg. A step-by-step guide, asking for interests,...)
2. Extensive functionalities that allow the end-user to provide feedback on specific content pieces



- (eg. Comments section, suggest edits,...)
- 3. A contributor community, with specific guidelines for contribution and curation
 - (eg. Process, do's & don'ts,...)
- 4. Workspace functionalities for end-users
 - (eg. Keeping notes, adding 'own' links to other resources,...)
- 5. Automated newsletter functionalities, that collects & combines relevant new content
 - (eg. Recommendations for you based on your search queries,...)
- 6. Communities of end-users that can connect on specific themes, topics or events
 - (eg. Share experience with other farmers, links to farmers who have used the solution, suggestions of 'friends',...)

5.2.4 Future

Stepping into the future, Farmbook could evolve towards a suite of modules & features that could enable thematic networks – and any type of multi-actor project – in their working.

A broader set of features could include:

1. A CMS (Content Management System) with website templates dedicated to the needs of multi-actor projects (*cfr. Squarespace, Wordpress,...*)
2. Built-in features for communication & dissemination
 - A design toolkit with visual identity components
 - Newsletter builder with templates (*cfr. Mailchimp,...*)
3. Built-in features for thematic networks to interact with their target audience and gauge for content needs
 - Learning Spaces & Webinar functionalities
 - Communities organized & managed by Thematic Networks

5.3 General usability and accessibility guidelines for the e-KRP

In the previous sections we have demonstrated that the creation of the database model for a HIKR benefits from having knowledge about the information needs of an end user in order to create a suitable ontology which allows the end user to find information using categories that fit the mental model of the end user when searching for information about the relevant topic. When bringing information from the back-end of the HIKR (i.e. the database) to the front-end (the digital platform with which the end user will interact), general knowledge about how an end user typically interacts with digital environments needs to be taken into account in order to create an environment that end users will use and keep on using.

Over the last two decades there has been a gradual change from perceiving digital products as mere information providers to digital products as more personalised experience providers which take into account the needs of relevant end users. Additionally, an EU directive (2016/2102) aimed at providing better access to digital products of public services for people with diverse abilities is currently shaping

the way digital products are being developed by forcing digital designers to take into account the point of view of different end users. In this section we'll provide an overview of guidelines to take into account when designing digital products to comply with usability and accessibility standards expected by end users.

5.3.1 Recommendations on usability & accessibility

We have collected the state-of-the art in usability & accessibility guidelines. These can be found in ANNEX 3.

Information presentation

The platform and its interface should consistently provide information in terminology that is understood and expected by as many relevant end users as possible.

The provided information needs to be concise and to the point as we want to prevent that end users are overwhelmed and dissuaded from using the platform due to an overload of information.

The most important information needs to be presented first and clearly marked so end users notice it first and can decide whether the information is relevant to their search.

The information flow

End users need to be provided with appropriate information as to where they are on the platform and how they can get back to a reference point such as the homepage.

A central navigation needs to be accessible on each page of the platform to allow end users to navigate through the platform at all times.

The central navigation should offer a limited amount (between 4 and 6) of clearly defined categories to choose from in order to prevent end users from doubting which category they need to choose to attain the correct information.

A search bar should be easily accessible to accommodate end users who choose to perform a specific search query.

Visual design

The information needs to be presented in an attractive way in line with contemporary standards of attractive visual design. The visual design cannot be outdated and needs to fit the expectations of the end users as this could give end users the impression that the presented information is outdated as well.

6. Conclusions & Next Steps

In task 3.2, we have:

- Developed and validated a data model which seems to be future-proof. This data model will serve as the basis for the EURAKNOS platform (e-KRP) and will be presented as a guideline to current & future Thematic Networks.
- Decided on a technology stack that is in line with multiple other projects that are either initiated or funded by the European Commission. This to ensure maximum interoperability with other & future projects, and to allow other teams to continue to work on the efforts in the EURAKNOS project
- Suggested a potential phasing of features for the EURAKNOS platform, beyond the scope of the EURAKNOS project.
- Researched the state-of-the-art in usability & accessibility guidelines to enable a better design of the user experience & user interface. This will ultimately ensure a better adoption of the platform by the end-users: farmers, foresters & advisors.

The outputs of D3.2 will feed into the technical guidelines for a High-Impact Knowledge Reservoir (HIKR) (D3.6), and into WP4 for the development of the prototype of the EURAKNOS Platform.



7. Annex 1: Classes of the EURAKNOS ontology and description of their properties

The EURAKNOS data model is provided in the form of an ontology consisting of classes, each of which possesses a number of properties (datatype properties and object properties). The identification and definition of class properties has drawn upon **schema.org**, which is a “*collaborative, community activity with a mission to create, maintain, and promote schemas for structured data on the Internet, on web pages, in email messages, and beyond.*”⁵³ Schema.org provides a **vocabulary** for defining entities, relationships between entities and actions. The schema.org vocabulary includes **types** relevant to the purpose of defining the classes of the EURAKNOS ontology and their properties. These types are: AudioObject, DigitalDocument, ImageObject, PresentationDigitalDocument, VideoObject, SoftwareApplication, CreativeWork, Person, and Organization. Table 4 provides an alignment between the EURAKNOS ontology’s classes and types of the schema.org vocabulary considered.

Table 4: Alignment between classes of the EURAKNOS ontology (data model) and types of the schema.org vocabulary

| Class and subclasses of the EURAKNOS ontology | | Schema.org vocabulary type | Link to schema.org vocabulary |
|---|--------------|-----------------------------|---|
| Output | Audio | AudioObject | https://schema.org/AudioObject |
| | Dataset | Dataset | https://schema.org/Dataset |
| | Document | DigitalDocument | https://schema.org/DigitalDocument |
| | Image | ImageObject | https://schema.org/ImageObject |
| | Presentation | PresentationDigitalDocument | https://schema.org/PresentationDigitalDocument |
| | Software | SoftwareApplication | https://schema.org/SoftwareApplication |
| | Video | VideoObject | https://schema.org/VideoObject |
| Creator | Individual | Person | https://schema.org/Person |
| | Organisation | Organization | https://schema.org/Organization |
| ThematicNetwork | | Thing | https://schema.org/Thing |
| Domain | | | |
| Topic | | | |
| EndUserType | | Person | https://schema.org/Person |

Schema.org applies a hierarchical structure for the arrangement of types, which means that, through the adoption of a supertype – subtype approach, properties can be inherited from the parent type to its subtypes and sub-subtypes. The most generic, top-level type of the schema.org vocabulary is *Thing*; in other words, a type used to represent any kind of entity/concept/object. All the other types are subtypes directly related to the *Thing* type, subtypes of direct subtypes of the *Thing* type and so on. The schema.org vocabulary types that have been considered for identifying properties for the classes of EURAKNOS ontology are part of this hierarchical structure as shown in Figure 11 below.

As made evident by the schema.org vocabulary’s hierarchical structure, and by taking account of the inheritance concept implied by this structure, a type inherits properties from its parent types whether linked in a direct or indirect way to them. For example, the *VideoObject* type is a subtype of the *MediaObject* type, which in turn is subtype of the *CreativeWork* type. So, the *VideoObject* type inherits all the properties of the *MediaObject* and *CreativeWork* types, but it also has a number of properties of its own. The fact that *Thing* is the most generic type in the hierarchy means

⁵³ <https://schema.org/>

that all other vocabulary types inherit the `Thing` type's properties. Therefore, the `VideoObject` type inherits the properties of the `Thing` type as well. The properties selected for each of the ontology classes are not solely properties of the respective schema.org types, but also properties of their parent types. The `Audio` class of the EURAKNOS ontology, for example, has been assigned a set of properties derived from the property list of the `AudioObject` type, as well as from the property lists of the parent types of `AudioObject` (namely, `MediaObject`, `CreativeWork` and `Thing`).

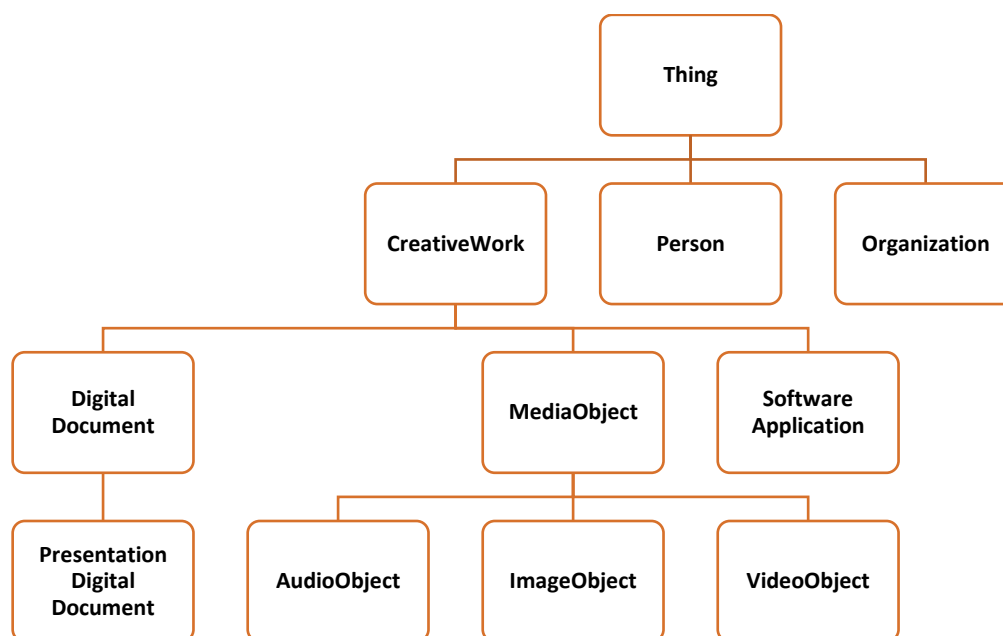


Figure 11: Schema.org types, in the context of the hierarchical structure employed by the schema.org vocabulary, used for identifying and defining properties for the EURAKNOS ontology's classes

It needs to be noted that the schema.org vocabulary does not provide any types that could directly be associated with the `ThematicNetwork`, `Domain`, and `Topic` ontology classes. For this reason, properties of the more generic, top-level `Thing` type have been considered for the needs of property identification for these ontology classes. In addition to that, apart from drawing upon the schema.org vocabulary for identifying properties for the classes of the EURAKNOS ontology, a number of custom properties have been defined per class to effectively address the particularities of the entities needed to be modelled.

Tables 5 to 10 provide the names and descriptions of properties per ontology class together with their types (i.e. datatype or object properties) and values.

Table 5: Properties of the `ThematicNetwork` class of the EURAKNOS ontology

| property | description | type | values |
|-------------|--|-------------------|--------|
| acronym | The acronym of the Thematic Network. | Datatype property | String |
| fullName | The full name of the Thematic Network. | Datatype property | String |
| url | URL of the Thematic Network's website. | Datatype property | String |
| description | Textual description of what the Thematic Network is about. | Datatype property | String |

| | | | |
|---------------|--|-----------------|-------------------------------------|
| relatesTo | Relationship between the ThematicNetwork class and the Domain class. i.e. a Thematic Network relates to one or more Domains. | Object property | Instances of the Domain class |
| CoordinatedBy | Relationship between the ThematicNetwork class and the Organisation class. i.e. One or zero Thematic Networks are coordinated by an Organisation. | Object property | Instances of the Organisation class |

Table 6: Properties of the Domain class of the EURAKNOS ontology

| property | description | type | values |
|-------------|---|-------------------------------|--|
| name | The name of the Thematic Network's domain. | Object property ⁵⁴ | Crop farming practices, actors, systems, types of crops, infrastructures, tools, technologies, climatic zones, environment, socio-economic context Animal husbandry and welfare practices, actors, systems, types of animals, infrastructures, tools, technologies, climatic zones, environment, socio-economic context Fisheries and aquaculture practices, actors, systems, types of fish, infrastructures, tools, technologies, climatic zones, environment, socio-economic context Forestry practices, actors, systems, types of forests, infrastructures, tools, technologies, climatic zones, environment, socio-economic context |
| description | Textual description of the Thematic Network's domain. | Datatype property | String |

Table 7: Properties of the Topic class of the EURAKNOS ontology

| property | description | type | values |
|----------|---|-------------------|--------|
| name | The name of the topic addressed by a specific output of a Thematic Network. | Datatype property | String |

⁵⁴ In some popular ontology editors like Protégé (<https://protege.stanford.edu/>), properties that receive a fixed number of values are modelled as object properties. This means that when modelling these properties, the modeller has to define an ontology class whose values are the property values.

| | | | |
|----------------|--|-------------------|--|
| description | Textual description of the output's topic. | Datatype property | String |
| associatedWith | Relationship between the <code>Topic</code> class and the <code>Domain</code> class. | Object property | Instances of the <code>Domain</code> class |

Table 8: Properties of the *Output* class of the EURAKNOS ontology

| property | description | type | values |
|-----------------------|--|-------------------|--|
| title | The title of an output of a Thematic Network | Datatype property | String |
| titleInLanguage | The language(s) in which the title of an output is made available. | Object property | {Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Irish, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovene, Spanish, Swedish} |
| description | Textual description of what an output is about. | Datatype property | String |
| descriptionInLanguage | The language(s) in which the textual description of an output is available. | Object property | {Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Irish, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovene, Spanish, Swedish} |
| contentInLanguage | The language(s) in which the content of an output is available. | Object property | {Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Irish, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovene, Spanish, Swedish} |
| keywords | A list of keywords that are associated with an output of the Thematic Network. | Datatype property | String |
| keywordsInLanguage | The language(s) in which an output-related keywords are made available. | Object property | {Bulgarian, Croatian, Czech, Danish, Dutch, English, Estonian, Finnish, French, German, Greek, Hungarian, Irish, Italian, Latvian, Lithuanian, Maltese, Polish, Portuguese, Romanian, Slovak, Slovene, Spanish, Swedish} |

| property | description | type | values |
|--------------------|--|-------------------|--|
| | | | Slovak, Slovene, Spanish, Swedish} |
| url | URL to the “source” of the output (where applicable). | Datatype property | String |
| comment | Comments from users of the output that may be probably available. | Datatype property | String |
| commentCount | Number of comments made to an output of a Thematic Network (if comments from users are available). | Datatype property | Integer |
| aggregateRating | The overall rating that has been made to an output (if user ratings are available). | Datatype property | Single-precision floating-point number ⁵⁵ |
| dateCreated | The date on which an output was created (full date or year). | Datatype property | String |
| accessibleForFree | A flag to signal whether an output is accessible for free or not. | Datatype property | Boolean |
| license | A description of specific licensing terms (e.g. a Creative Commons license) that describe how an output may be further exploited. | Datatype property | String |
| geographicLocation | The geographic location (country/-ies) to which the content of an output (for instance, solution to a problem or innovative practice) relates. | Object property | {Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom} |
| version | The version of a Thematic network’s output that is available. | Datatype property | String |
| purpose | The reason/purpose for which an output has been produced. | Object property | {access to data, communication, dissemination, education/training} |
| fileSize | The size of the file (in KBs/MBs) through which an output of a Thematic Network is available. | Datatype property | String |

⁵⁵ According to Wikipedia (https://en.wikipedia.org/wiki/Single-precision_floating-point_format), a single-precision floating-point number format is a computer number format, usually occupying 32 bits in computer memory.

| property | description | type | values |
|-----------|--|-----------------|------------------------------------|
| hasTopic | Relationship between the Output class and the Topic class. | Object property | Instances of the Topic class |
| createdBy | Relationship between the Output class and the Creator class. | Object property | Instances of the Creator class |
| usedBy | Relationship between the Output class and the EndUserType class. | Object property | Instances of the EndUserType class |

Table 9: Properties of the EndUserType class of the EURAKNOS ontology

| property | description | type | values |
|-------------|---|-------------------|---|
| name | The name of an end-user type that is going to use an output of a Thematic Network. | Object property | {advisor, bee keeper, consumer, distributor, entrepreneur, farmer, fisherman, forester, instructor, livestock breeder, policy maker, researcher, student, trader, educator/trainer, university professor} |
| description | Description of an end-user type that is going to use an output of a Thematic Network. | Datatype property | String |

Table 10: Properties of the Creator class of the EURAKNOS ontology

| property | description | type | values |
|-------------|---|-------------------|--|
| name | The name of an output creator. | Datatype property | String |
| description | Textual description of an output creator. | Datatype property | String |
| country | The country of origin of an output creator. | Object property | {Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom} |

As made evident in the EURAKNOS ontology (i.e. data model), the Audio, Dataset, Document, Image, Presentation, Software and Video classes are subclasses of the Output class, which means that all the properties of the Output class are inherited by its subclasses. Yet, the subclasses of the Output class have also a number of additional properties. These additional properties are presented in detail in Tables 11 to 17 below.

Table 11: Properties of the Audio class of the EURAKNOS ontology

| property | description | type | values |
|----------|-------------|------|--------|
|----------|-------------|------|--------|

| | | | |
|------------|---|-------------------|--|
| duration | The duration of an audio object. | Datatype property | String |
| format | The file format of an audio object (available value options adhere to open file formats). | Object property | {.m4a, .caf, .flac, .mpc, .mp+, .mpp, .mp3, .bit, .ogg, .ogv, .oga, .ogx, .ogm, .spx, .opus, .spx, .wv} |
| playerType | The player type required for playing an audio file. | Datatype property | String |
| transcript | The transcript of an audio object. | Datatype property | String |
| type | The type of an audio object. | Object property | {educational/training podcast, event capturing podcast, interview, solo podcast, tutorial, guide, commentary, panel discussion, question-and-answer podcast, on-demand seminar, audio magazine, other} |

Table 12: Properties of the *Dataset* class of the EURAKNOS ontology

| property | description | type | values |
|--------------------|--|-------------------|---|
| format | The file format of a dataset. | Datatype property | String |
| issn | The International Standard Serial Number (ISSN) that identifies a dataset that has been published. | Datatype property | String |
| measumentTechnique | A technique or technology used in a dataset, corresponding to the method used for measuring the variable(s) (described using “ <i>variableMeasured</i> ”). This is oriented towards scientific and scholarly dataset publication but may have broader applicability; it is not intended as a full representation of measurement, but rather as a high-level summary for dataset discovery. | Datatype property | String |
| type | The type of a dataset. | Object property | {auditory data, crop-related data, geospatial data, graph-related data, imagery data, input-related data, network-related data, temporal data, textual data, video data, yield-related data, other} |
| variableMeasured | A property that indicates the variables measured in a dataset. | Datatype property | String |

Table 13: Properties of the *Document* class of the EURAKNOS ontology

| property | description | type | values |
|----------|--|-----------------|--|
| format | The file format of a document (available value options adhere to open file formats). | Object property | {.txt, .csv, .htm/.html, markdown, .dvi, .dbk, .xml, .epub, .fb2, .fbz, .docx, .docm, .doc, .odt, .fodt, .oxps, .xps, .pdf, .ps, .xls, .xlsx, .xhtml, .xht} |
| type | The type of a document. | Object property | {article in conference proceedings, book, booklet, brochure/flyer, chapter in edited volume, deliverable report, factsheet, handbook, guide, journal article, manual, milestone report, newsletter, practice abstract, press release, spreadsheet, review document, report/paper, technical article, technical information/specifications card, tutorial, other} |

Table 14: Properties of the *Image* class of the EURAKNOS ontology

| property | description | type | values |
|----------|--|-------------------|--|
| height | The height of an image object. | Datatype property | Integer |
| width | The width of an image object. | Datatype property | Integer |
| format | The file format of the image object (available value options adhere to open file formats). | Object property | {.png, .apng, .flif, .gbr, .gif, .jp2, .j2k, .jpf, .jpx, .jpm, .mj2, .mng, .exr, .pdf, .png, .svg, .svgz, .webp, .xpm} |
| type | The type of the image object. | Object property | {chart/graph, static figure/image, interactive figure/image, infographic, static map, interactive map, other} |

Table 15: Properties of the *Presentation* class of the EURAKNOS ontology

| property | description | type | values |
|----------|--|-----------------|--|
| format | The file format of a presentation (available value options adhere to open file formats). | Object property | {.pdf, .potm, .potx, .ppt, .pptx} |
| type | The type of a presentation. | Object property | {educational/training presentation, guide, informative presentation, tutorial, motivational presentation, decision-making presentation, problem-solving presentation, other} |

Table 16: Properties of the *Software* class of the EURAKNOS ontology

| property | description | type | values |
|-------------------|--|-------------------|--------|
| availableOnDevice | The device (or type of device) required to run an application. Used in cases where a specific make/model is required to run the application. | Datatype property | String |

| | | | |
|----------------------|--|-------------------|--|
| format | The file format of a software application. | Object property | {.exe, .app, .vb, .scr} |
| memoryRequirements | Minimum memory requirements. | Datatype property | String |
| operatingSystem | Operating systems supported. | Datatype property | String |
| softwareRequirements | Component dependency requirements for application. This includes runtime environments and shared libraries that are not included in the application distribution package, but required to run the application (Examples: DirectX, Java or .NET runtime). | Datatype property | String |
| storageRequirements | Free disk space required to install an application to a computer. | Datatype property | String |
| type | The type of a software application. | Object property | {AI software, business software, data repository/database, data analysis software, decision support tool, educational/training software, FMIS, game, scientific software, simulation, other} |

Table 17: Properties of the *Video* class of the EURAKNOS ontology

| property | description | type | values |
|------------|--|-------------------|--|
| duration | The duration of a video object. | Datatype property | String |
| format | The file format of a video (available value options adhere to open file formats).. | Object property | {.drc, .mkv, .mk3d, .mka, .mks, .mp4, .m4a, .m4p, .m4b, .m4r, .m4v, .webm, .ogv, .ogg} |
| playerType | The player type required to play a video file. | Datatype property | String |
| transcript | The transcript of a video object. | Datatype property | String |
| type | The type of a video object. | Object property | {vlog, webinar, guide, event capturing video, tutorial/how-to video, presentation/live talk capturing video, interview video, product/ feature review video, documentary video, testimonial, question-and-answer video, case study, simulation video, educational/training video, event capturing video, informative video, other} |

| | | | |
|----------------|------------------------------|-------------------|--------|
| videoFrameSize | The frame size of the video. | Datatype property | String |
| videoQuality | The quality of the video. | Datatype property | String |

The `Individual` and `Organisation` classes are subclasses of the `Creator` class meaning that all the properties of the `Creator` class are inherited by all its subclasses. However, the subclasses of the `Creator` class have a number of additional properties as well. These additional properties are presented in detail in Tables 18 and 19 below.

Table 18: Properties of the *Individual* class of the EURAKNOS ontology

| property | description | type | values |
|----------------|---|-------------------|--|
| email | The e-mail address of an individual. | Datatype property | String |
| type | The type of an individual. | Object property | {adviser, bee keeper, consumer, distributor, entrepreneur, farmer, fisherman, forester, instructor, livestock breeder, policy maker, researcher, student, trader, educator/trainer, university professor, other} |
| affiliatedWith | Relationship between the Individual subclass and the Organisation subclass. | Object property | Instances of the Organisation class |

Table 19: Properties of the *Organisation* class of the EURAKNOS ontology

| property | description | type | values |
|--------------|---|-------------------|--|
| areaServed | The geographic area where a service or item is provided by an organisation. | Object property | {Austria, Belgium, Bulgaria, Croatia, Cyprus, Czech Republic, Denmark, Estonia, Finland, France, Germany, Greece, Hungary, Ireland, Italy, Latvia, Lithuania, Luxembourg, Malta, Netherlands, Poland, Portugal, Romania, Slovakia, Slovenia, Spain, Sweden, Switzerland, United Kingdom} |
| contactPoint | Contact point for an organization. | Datatype property | String |
| involvement | Type of involvement an organization has in the Thematic Network. | Object property | {leading organization, participating organization} |
| type | The type of an organisation. | Object property | {advisors' association, AKIS community, consulting company, consumers' association, distribution company, environmental organisation, farmers' association, food industry, input industry, NGO, public authority, retailer, research institute, school, SME, technical institute, trading company, training organization, university, other} |

8. Annex 2: Details about the advisor survey conducted in the context of the EUFRAS meeting in Athens

The survey results presented in the “**Considerations for selecting data objects to store into the High Impact Knowledge Reservoir**” section have been obtained from an advisor survey that took place in the context of the EUFRAS meeting in Athens (24 - 25 February, 2020). The survey questionnaire that was used for the needs of the survey was designed by PSKW and comprised a total number of 12 items. The full list of questionnaire items is provided below. The sample of survey participants consisted had a size of 32 ($n = 32$). As made evident from the graphs presented in Figures 12 to 14 below, there was an equal representation of females and males in the sample (female advisors = 43.8%; male advisors = 50.0%).

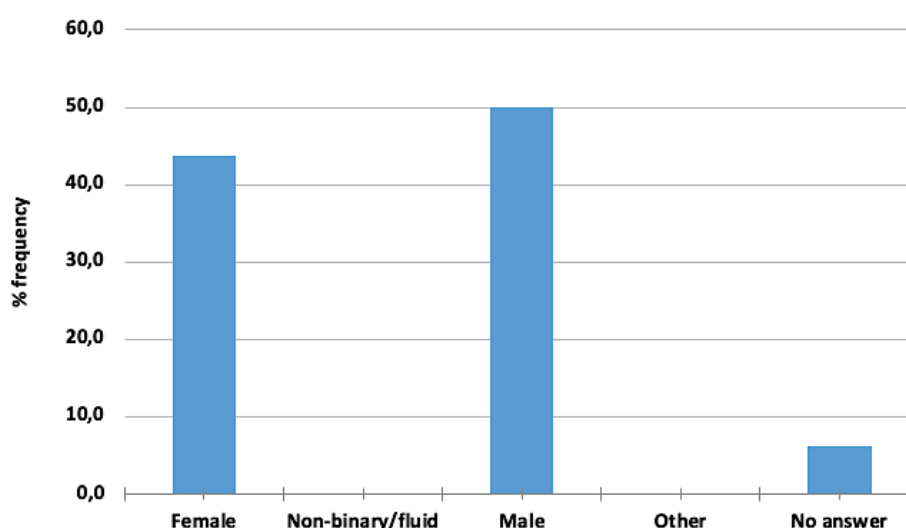


Figure 12: Gender distribution in the sample of advisors that participated in the EUFRAS meeting survey

The majority of sample subjects belong to the 40-50 and the 50-60 age groups with the frequencies of these age groups being 29.0% and 25.8% respectively. The frequency of subjects in the age range between 30 and 40 years old was 16.1%, whereas 12.9% of the sample subjects have an age between 20 and 30 years old. The age group that was least represented in the sample was that from 60 to 70 years old, which accounted for 9.7% of the sample size.

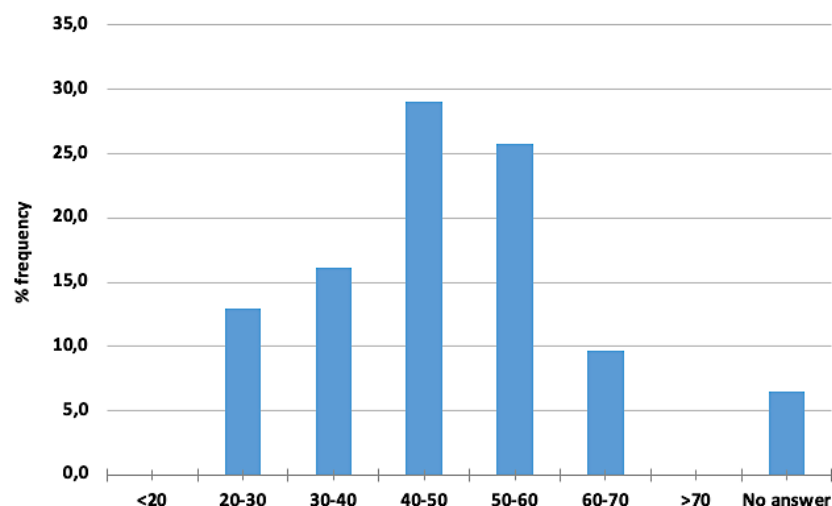


Figure 13: Distribution of ages in the sample of advisors that participated in the EUFRA meeting survey

As far as nationalities are concerned, the majority of survey participants come from Greece (15.6%), 12.5% are Bulgarian, 9.4 % are German, and then we have Latvian, Estonian, Spanish participants, as well as participants from Switzerland (6.3%). The rest of the nationalities were represented at a rate of 3.1%.

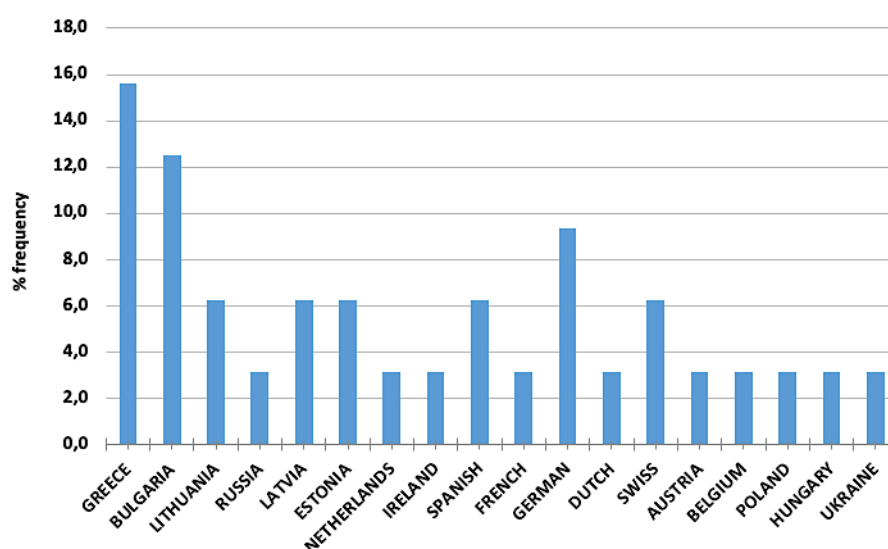


Figure 14: Nationalities of survey participants

Q1: There are new project results available. How would you like to be informed about this?
(multiple choice)

1. Technical magazines
2. Websites of research stations
3. Newsletters
4. Meetings, Study days
5. Facebook
6. WhatsApp
7. Twitter
8. LinkedIn
98. Other (Specify:)
99. No answer

Q2: Where do you search for the information you need? (multiple choice)

1. Technical magazines
2. Websites of research stations
3. Search engines
4. Scientific databases
98. Other (Specify:)
99. No answer

Q3: In what language do you search for information? (multiple choice)

1. English
2. Native language
98. Other (Specify:)
99. No answer

Q4: What kind of content/information would you prefer? (multiple choice)

1. Presentation of best practices
2. Presentation of innovative practices
3. Information material (just to know what is going on in the community)
4. Educational/training material
5. Guidelines
98. Other (Specify:)
99. No answer

Q5: Would you opt for being provided content/information based on your profile and history of interactions? (single answer)

1. Yes
2. No
99. No answer

Q6: In what phase of a project would you like to be involved? (multiple choice)

1. Project ideas based on your own experiences
2. Project writing
3. Participation in steering group/user group
4. Expert panel following up tests
5. Testing technologies
6. Participation in workshops, study days, meetings
7. Feedback on project impact
98. Other (Specify:)
99. No answer

Q7: What group do you belong to? (single answer)

1. Grower
2. Advisor
3. Researcher



4. Industry
5. Student
6. Teacher
7. Policy maker
98. Other (Specify:)
99. No answer

Q8: What is your nationality? (single answer)

98. Specify:
99. No answer

Q9: Which age category do you belong to? (single answer)

1. <20
2. 20 – 30
3. 30 – 40
4. 40 – 50
5. 50 – 60
6. 60 – 70
7. >70
99. No answer

Q10: What gender are you? (single answer)

1. Female
2. Non-binary/fluid
3. Male
98. Other (Specify:)
99. No answer

Q11: What kind of information would you like to be delivered as part of search results display?
(multiple choice)

1. Title of the provided result item.
2. Short textual description of the provided result item.
3. List of keywords associated with the provided result item.
4. Descriptive icon of the provided result item.
5. Format of the provided result item (pdf, ppt, jpg, mp4, etc.)
6. Type of provided result item (factsheet, leaflet, report, video, podcast, scientific article etc.)
98. Other (Specify:)
99. No answer

Q12: Which of the following kinds of digital content is most preferable to you? (multiple choice)

1. Audio
2. Image
3. Presentation (slideshow)
4. Text
5. Video
98. Other (Specify:)
99. No answer



9. Annex 3: State-of-the Art in Usability & Accessibility Guidelines

9.1 Usability guidelines

User experience design (UX-design) attempts to make the use of (digital) products as user friendly as is possible. In general, good UX-design will not be noticed by the user, while bad UX-design will. At a more abstract level UX-design tends to follow a set of guidelines which the designer tries to respect when designing a digital product. As digital technology can change rapidly, the implementation of these guidelines will differ depending on the available technology but also on trends set by dominant digital product providers (e.g. Apple, Facebook, Google). Instead of advising designers to use, for instance, a specific button and location to share content, designers are advised to detect and follow the current trend in order to make the experience over different digital products as seamless as possible. This also implies that a sustainable digital product is one which is updated regularly to match the usability trends prominent at that moment. Listed below are several best-practice lists of guidelines which show overlap, but help UX-designers create user friendly digital products.

1) TEN USABILITY HEURISTICS FOR INTERFACE DESIGN

In 1990 Jakob Nielsen and Rolf Molich⁵⁶ created a list of usability heuristics for user interface design which are still considered relevant and have been further refined through studies.

1. Visibility of system status

The system should always keep users informed about what is going on, through appropriate feedback within reasonable time.

2. Match between system and the real world

The system should speak the users' language, with words, phrases and concepts familiar to the user, rather than system-oriented terms. Follow real-world conventions, making information appear in a natural and logical order.

3. User control and freedom

Users often choose system functions by mistake and will need a clearly marked "emergency exit" to leave the unwanted state without having to go through an extended dialogue. Support undo and redo.

4. Consistency and standards

Users should not have to wonder whether different words, situations, or actions mean the same thing.

5. Error prevention

Even better than good error messages is a careful design which prevents a problem from occurring in the first place. Either eliminate error-prone conditions or check for them and present users with a confirmation option before they commit to the action.

6. Recognition rather than recall

Minimize the user's memory load by making objects, actions, and options visible. The user should not have to remember information from one part of the dialogue to another. Instructions for use of the system should be visible or easily retrievable whenever appropriate.

⁵⁶ <https://www.nngroup.com/articles/ten-usability-heuristics/>

7. Flexibility and efficiency of use

Accelerators — unseen by the novice user — may often speed up the interaction for the expert user such that the system can cater to both inexperienced and experienced users. Allow users to tailor frequent actions.

8. Aesthetic and minimalist design

Dialogues should not contain information which is irrelevant or rarely needed. Every extra unit of information in a dialogue competes with the relevant units of information and diminishes their relative visibility.

9. Help users recognize, diagnose, and recover from errors

Error messages should be expressed in plain language (no codes), precisely indicate the problem, and constructively suggest a solution.

10. Help and documentation

Even though it is better if the system can be used without documentation, it may be necessary to provide help and documentation. Any such information should be easy to search, focused on the user's task, list concrete steps to be carried out, and not be too large.

2) TEN PRINCIPLES FOR GOOD DESIGN

In the 1960 Dieter Rams⁵⁷ defined Ten principles for good design which apply to design in general but have found their way to the minds of many digital designers.

1. Good design is innovative

The possibilities for innovation are not, by any means, exhausted. Technological development is always offering new opportunities for innovative design. But innovative design always develops in tandem with innovative technology, and can never be an end in itself.

2. Good design makes a product useful

A product is bought to be used. It has to satisfy certain criteria, not only functional, but also psychological and aesthetic. Good design emphasises the usefulness of a product whilst disregarding anything that could possibly detract from it.

3. Good design is aesthetic

The aesthetic quality of a product is integral to its usefulness because products we use every day affect our person and our well-being. But only well-executed objects can be beautiful.

4. Good design makes a product understandable

It clarifies the products structure. Better still, it can make the product talk. At best, it is self-explanatory.

5. Good design is unobtrusive

Products fulfilling a purpose are like tools. They are neither decorative objects nor works of art. Their design should therefore be both neutral and restrained, to leave room for the users self-expression.

6. Good design is honest

⁵⁷ <https://www.designprinciplesftw.com/collections/ten-principles-for-good-design>

It does not make a product more innovative, powerful or valuable than it really is. It does not attempt to manipulate the consumer with promises that cannot be kept.

7. Good design is long-lasting

It avoids being fashionable and therefore never appears antiquated. Unlike fashionable design, it lasts many years - even in today's throwaway society.

8. Good design is thorough down to the last detail

Nothing must be arbitrary or left to chance. Care and accuracy in the design process show respect towards the consumer.

9. Good design is environmentally friendly

Design makes an important contribution to the preservation of the environment. It conserves resources and minimises physical and visual pollution throughout the lifecycle of the product.

10. Good design is as little design as possible

Less, but better - because it concentrates on the essential aspects, and the products are not burdened with non-essentials.

Back to purity, back to simplicity.

3) EIGHT GOLDEN RULES FOR INTERFACE DESIGN

In 1987 Ben Shneiderman⁵⁸ distilled Eight Golden Rules of Interface Design from the research he did in Human Computer Interaction.

1. Strive for consistency

Consistent sequences of actions should be required in similar situations; identical terminology should be used in prompts, menus, and help screens; and consistent colour, layout, capitalization, fonts, and so on, should be employed throughout. Exceptions, such as required confirmation of the delete command or no echoing of passwords, should be comprehensible and limited in number.

2. Seek universal usability

Recognize the needs of diverse users and design for plasticity, facilitating transformation of content. Novice to expert differences, age ranges, disabilities, international variations, and technological diversity each enrich the spectrum of requirements that guides design. Adding features for novices, such as explanations, and features for experts, such as shortcuts and faster pacing, enriches the interface design and improves perceived quality.

3. Offer informative feedback

For every user action, there should be an interface feedback. For frequent and minor actions, the response can be modest, whereas for infrequent and major actions, the response should be more substantial. Visual presentation of the objects of interest provides a convenient environment for showing changes explicitly (see the discussion of direct manipulation in Chapter 7).

4. Design dialogs to yield closure

Sequences of actions should be organized into groups with a beginning, middle, and end. Informative feedback at the completion of a group of actions gives users the satisfaction of accomplishment, a sense of relief, a signal to drop contingency plans from their minds, and an indicator to prepare for the

⁵⁸ <https://www.cs.umd.edu/~ben/goldenrules.html>

next group of actions. For example, e-commerce websites move users from selecting products to the checkout, ending with a clear confirmation page that completes the transaction.

5. Prevent errors

As much as possible, design the interface so that users cannot make serious errors; for example, gray out menu items that are not appropriate and do not allow alphabetic characters in numeric entry fields (Section 3.3.5). If users make an error, the interface should offer simple, constructive, and specific instructions for recovery. For example, users should not have to retype an entire name-address form if they enter an invalid zip code but rather should be guided to repair only the faulty part. Erroneous actions should leave the interface state unchanged, or the interface should give instructions about restoring the state.

6. Permit easy reversal of actions

As much as possible, actions should be reversible. This feature relieves anxiety, since users know that errors can be undone, and encourages exploration of unfamiliar options. The units of reversibility may be a single action, a data-entry task, or a complete group of actions, such as entry of a name-address block.

7. Keep users in control

Experienced users strongly desire the sense that they are in charge of the interface and that the interface responds to their actions. They don't want surprises or changes in familiar behaviour, and they are annoyed by tedious data-entry sequences, difficulty in obtaining necessary information, and inability to produce their desired result.

8. Reduce short-term memory load

Humans' limited capacity for information processing in short-term memory (the rule of thumb is that people can remember "seven plus or minus two chunks" of information) requires that designers avoid interfaces in which users must remember information from one display and then use that information on another display. It means that cellphones should not require re-entry of phone numbers, website locations should remain visible, and lengthy forms should be compacted to fit a single display.

4) DESIGN PRINCIPLES FOR REDUCING COGNITIVE LOAD

More recently Jon Yablonski⁵⁹ introduced several design principles to reduce cognitive load which prove to be helpful as an overload of information from different sources is making it increasingly difficult for users to process information provided without taking these principles into account.

1. Avoid Unnecessary Elements

Like everything in design, less is more. Any element that isn't helping the user achieve their goal is working against them because they must process it and store it in working memory, alongside the things that will help them. Avoiding excessive colours, imagery, design flourishes, or layouts that don't add value is crucial. But simplicity comes with a caveat: don't overvalue it at the cost of clarity.

2. Leverage Common Design Patterns

By leveraging common design patterns when it makes sense, you are giving the user familiar elements which they already understand. This in turn reduces the amount of learning they need to do, thus

⁵⁹ <https://jonyablonski.com/articles/2015/design-principles-for-reducing-cognitive-load/>

enabling them to move right along and get closer to achieving their goal. My favourite sources for design pattern inspiration are Design Patterns on CodePen and the Blueprint Archives on Codrops.

3. Eliminate Unnecessary Tasks

Anywhere you are asking the user to read content, remember information or make a decision contributes to cognitive load. Whenever possible, it is good to shift these tasks away from the user and make it easier for them to stay focused on their goal. While it isn't possible to remove all tasks, there is usually an opportunity to offload some task by setting defaults that can be edited, or leveraging previously entered information. Some companies are even taking this a step further with anticipatory design.

4. Minimize Choices

As previously mentioned, our working memory is limited. When confronted with too many choices, cognitive load will increase due to decision paralysis. It is important that we minimize the choices the user must make at any given moment, especially in places such as navigation, forms, and drop-downs.

5. Display Choices as a Group

When choices are split into separate groups and hidden, users often mistake the options that are visible as the complete group. This means that users are likely to never find the additional choices, which not only limits what is available to them, but also makes it more difficult to decide on which option to select because they are not aware of the alternatives. Therefore, it is best to eliminate the resulting cognitive load by always displaying choices as a group.

6. Strive for Readability

Making our content legible isn't enough — we need to make it readable. This means our typography must be aesthetically pleasing, appropriate for the content and easy to read while design remains relatively invisible. By doing this, we can ensure there are as little distractions as possible for the user, which results in a better understanding of the content by the user.

7. Use Iconography with Caution

Research has shown that iconography can be hard to memorize and, contrary to intuition, can increase cognitive load by requiring mental processing to infer meaning or recognize. While universally understood icons work well (ie. print, close, play/pause, reply, tweet, share on Facebook), most are subject to the user's understanding based on previous experience (in which there is no standard). When leveraging the power of iconography, it is best to accompany them with text labels to communicate the meaning and reduce ambiguity.

9.1.1 Accessibility principles and guidelines

The European Commission and Council aim to create a more digitally inclusive environment and have create a directive for digital products of public services.

The EU directive 2016/2102⁶⁰:

- covers websites and apps of public sector bodies, with a limited number of exceptions (e.g. broadcasters, live streaming);
- refers to specific standards to make websites and mobile apps more accessible. Such standards require for instance that there should be a text description for images, or that users are able to interact with a website without using a mouse, which can be difficult for some people with disabilities;

⁶⁰ <https://eur-lex.europa.eu/eli/dir/2016/2102/oj>

- requires the publication of an accessibility statement for each website and mobile app, describing the level of accessibility and indicating any content that is not accessible;
- calls for a feedback mechanism which the users can use to flag accessibility problems or to ask for the information contained in a non-accessible content;
- expects regular monitoring of public sector websites and apps by Member States, and that they report on the results of the monitoring. These reports have to be communicated to the Commission and to be made public for the first time by 23 December 2021.

The World Wide Web Consortium (W3C)⁶¹, involved in developing Web standards, has created a set of accessibility standards and guidelines which aid digital designers and developers to create accessible digital products.

Examples include:

- Providing text alternatives for non-text content such as images (through descriptions) and video or audio files (through transcripts)
- Allowing navigation through a digital product without having to rely on a mouse (easy navigation through keyboard)
- Users can easily navigate, find content, and determine where they are
- Content appears and operates in predictable ways by providing a easily identifiable navigation and using consistency in the labelling of the navigation categories
- Content is compatible with current and future user tools by providing a name, role, and value for non-standard user interface components

⁶¹<https://www.w3.org/standards/webdesign/accessibility>,
<https://www.w3.org/WAI/fundamentals/accessibility-principles/>

10. References

- Agricultural European Innovation Partnership** (2016). Thematic Networks under Horizon 2020: Compiling knowledge ready for practice. Retrieved from https://ec.europa.eu/eip/agriculture/sites/agri-eip/files/eip-agri_brochure_thematic_networks_2016_en_web.pdf
- Agricultural European Innovation Partnership** (2017). Horizon 2020 multi-actor projects. Retrieved from https://ec.europa.eu/eip/agriculture/sites/agri-eip/files/eip-agri_brochure_multi-actor_projects_2017_en_web.pdf
- Baclawski, K., et al.** (2001). Extending UML to support ontology engineering for the semantic web. In *International Conference on the Unified Modeling Language* (pp. 342-360). Springer, Berlin, Heidelberg.
- Booch, G., et al.** (2005). *The Unified Modeling Language User Guide*. Addison-Wesley Professional.
- Djurić, D., et al.** (2004). A UML profile for OWL ontologies. In *Model Driven Architecture* (pp. 204-219). Springer, Berlin, Heidelberg.
- Gernert, M., et al.** (2019). Report on added value and feasibility of e-KRP. EURAKNOS – Connecting Thematic Networks as Knowledge Reservoirs towards a European Agricultural Knowledge Innovation Open Source System Deliverable 4.1 – WP4.
- Marois, M., et al.** (2019). Report on format and design of knowledge reservoirs of Thematic Networks. EURAKNOS – Connecting Thematic Networks as Knowledge Reservoirs towards a European Agricultural Knowledge Innovation Open Source System Deliverable 2.2 – WP2.
- Marois, M., et al.** (2020). Report on HIKR methodologies and practices for data generation, selection, collection and storage. EURAKNOS – Connecting Thematic Networks as Knowledge Reservoirs towards a European Agricultural Knowledge Innovation Open Source System Deliverable 3.1 – WP1.
- Mosquera Losada, M. R., et al.** (2019). Report on content and structures of data provision and knowledge formats of TNs. EURAKNOS – Connecting Thematic Networks as Knowledge Reservoirs towards a European Agricultural Knowledge Innovation Open Source System Deliverable 2.1 – WP2.
- Wilkinson, M. D., et al.** (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 3.