



KINDRA DELIVERABLE D3.1

DRAFT SYNTHESIS OF COUNTRY REPORTS

Summary:

The present document summarizes the draft synthesis of data, which are included in the Country reports and the EIGR (European Inventory of Groundwater Research). The Country reports and most of the EIGR records are provided by the experts from the EFG's national member associations (Linked Third Parties). The experts gathered the information from diverse sources and used the terminology and the guidelines, created in WP1. This deliverable contains a qualitative and quantitative assessment of the resources uploaded by LTPs to EIGR using the developed HRC-SYS classification system and the barriers & gaps identified in the Country reports. The analytical tools for the data assessment (EIGR tools and VOS viewer) are also introduced in this deliverable. These tools will be used for more in depth analyses of groundwater research in following reports and papers e.g. when comparing EIGR data from the commercial research databases, Scopus and Web of Science. The metadata inserted into the EIGR until the end of February 2018 have been considered in the statistics.

Authors:

University of Miskolc: Viktória Mikita

Project acronym:	KINDRA
Project title:	Knowledge Inventory for hydrogeology research
Grant Agreement number:	642047
Call identifier:	H2020-WATER-2014-one-stage
Topic:	WATER-4a-2014
Start date of the project:	01/01/2015
Duration:	36 months
Website:	www.kindraproject.eu
This report has been produced with financial support from the European Union's Horizon 2020 research and innovation programme under grant agreement No 642047".	
The contents of this report are the sole responsibility of the KINDRA Consortium and can under no circumstances be regarded as reflecting the position of the European Union. The Executive Agency for Small and Medium-sized Enterprises (EASME) is not responsible for any use that may be made of the information it contains.	

Table of Contents

1	INTRODUCTION	5
2	OVERVIEW OF COUNTRY REPORTS	5
2.1	INFORMATION SOURCES	6
2.2	TYPE OF INFORMATION	10
2.3	TYPES OF DATA AND RESOURCES (TOPICS OF THE DATA)	10
2.4	GAPS AND BARRIERS.....	12
2.5	OVERVIEW OF EIGR METADATA	13
3	ANALYTICAL TOOLS FOR DATA ASSESSMENT	13
3.1	DEVELOPED EIGR TOOLS FOR DATA ANALYSIS AND VISUALISATION	13
3.2	EXTERNAL TOOLS	15
4	ASSESSMENT OF EIGR METADATA	17
4.1	DATA ASSESSMENT FOR HRC-SYS CLASSIFICATION.....	17
4.1.1	CO-OCCURRENCE ANALYSIS OF HRC-SYS CATEGORIES BY EIGR METADATA.....	20
4.1.2	THE KEYWORD OCCURRENCE BY COUNTRY.....	25
4.2	DATA ASSESSMENT FOR RESEARCH AND KNOWLEDGE CLASSES	30
4.3	DATA ASSESSMENT FOR TECHNOLOGICAL READINESS LEVEL CLASSES.....	31
5	SUMMARY AND CONCLUSIONS	32
	ANNEX 1.....	34

List of figures

Figure 1. Research and knowledge classes identified in the KINDRA project.....	10
Figure 2. The distribution of data types/resource categories in the EIGR entries reported in the country reports.....	12
Figure 3. The Total number of EIGR contents inserted by the LTPs and KINDRA consortium partners until 28/02/2018.....	13
Figure 4. The Keywords Cloud	14
Figure 5. The Resource Distribution Map.....	14
Figure 6. The Research Topics -Operational Actions 2D Chart (for each of the five Societal Challenges) of EIGR	15

Figure 7. Vosviewer heat map for co-occurrences of the most often used keywords for the Societal Challenge (SC) - Climate, Environment, and Resources, 1997-2016 (Mads Breum, 2016).....	16
Figure 8. Vosviewer network map for co-occurrences of the most often used keywords for the Societal Challenge (SC) - Climate, Environment, and Resources, 1997-2016 (Mads Breum, 2016).....	17
Figure 9. The HRC-SYS three overarching categories: 1) Societal Challenges, 2) Operational Actions and 3) Research Topics and their five main sub-categories (“keywords”) used for classification of the resources uploaded to EIGR.....	18
Figure 10. The distribution of EIGR metadata related to Societal Challenges (SC)	19
Figure 11. The distribution of EIGR metadata related to Operational Actions (OA)	19
Figure 12. The distribution of EIGR metadata related to Research Topics (RT)	20
Figure 13. The number of metadata for Societal challenges -Health and 5-5 classes of Research topics and Operational actions	21
Figure 14. The number of metadata for Societal challenges -Food and 5-5 classes of Research topics and Operational actions	22
Figure 15. The number of metadata in the intersections of Societal challenges -Energy and 5-5 classes of Research topics and Operational actions	23
Figure 16. The number of metadata in the intersections of Societal challenges -Climate, environment and resources and 5-5 classes of Research topics and Operational actions	24
Figure 17. The number of metadata in the intersections of Operational actions-Water supply and the 5-5 classes of Research topics and Societal challenges	25
Figure 18: The 10 most popular keywords for four countries	26
Errore. Il segnalibro non è definito.	
Figure 19: The keyword occurrences vs. the number of nations (Correira, 2018)	27
Figure 20: The most popular keywords in the EIGR (on the status 28/02/2018)	28
Figure 21. The distribution of EIGR metadata related to Research and knowledge classes (KC)	30
Figure 22. The distribution of EIGR metadata related to Technological Readiness level (TRL)	31

List of tables

Table 1. Summary of the Inventory data collection and the total number of entries by EFG LTPs until the end of 2017 (taken from D2.3)	6
Table 2. Summary of the most important sources of information per country	7
Table 3. Summary of the topics of data with the numbers of entries reported by the LTPs (source: D2.3)	11
Table 4. The 10 most popular keywords in the EIGR entries and their 4 frequent co-occurred keywords	29

1 INTRODUCTION

Within the framework of the WP2, the European Inventory of Groundwater Research (EIGR) has been developed to collect the Europe-wide hydrogeology related applied and scientific knowledge on international, national and regional levels. The EIGR makes the information available with open access for researchers and the public. The metadata insertion to the EIGR has been implemented by the European Federation of Geologists Linked Third Parties (LTP, 20 National Associations/Geological Societies) participating in the project. The Associations provided “Country reports” related to the qualitative status of the inserted EIGR metadata. The experts gathered the information from diverse sources and used the concept, terminology and the guidelines, created in WP1.

This deliverable follows the preliminary D2.3 (Country reports) & D2.4 (Datasheets), and contains the overview of Country reports and EIGR metadata. The qualitative/quantitative assessment of available information focused on the HRC-SYS, as well as barriers and gaps summarized in the Country reports. In this document the developed analytical tools of EIGR and an external tool -VOSviewer- are also presented.

A preliminary co-occurrence analysis has been done in this document and the metadata inserted into the EIGR until the end of February 2018 have been considered in the statistics.

2 OVERVIEW OF COUNTRY REPORTS

The EFG Linked Third Parties (LTPs) representing 21 European countries were asked to provide Country reports in 2016-2017.

The Country reports include 6 questions, including information about the overview of the uploaded metadata, time consumed for uploading to EIGR, the main source of information used during the data collection, classification of their information based on the “Research and knowledge classes”, classification of their entries into 7 categories (database, maps, books, etc.), barriers to find data and gaps in finding the information. The detailed description of the Country reports is in D2.3 (Country reports). The template of Country reports can be found in D2.3 (Country reports).

The EIGR metadata were inserted by the LTPs and the KINDRA consortium partners. There are 2200 published records in the EIGR, from this number 1986 metadata were uploaded by the LTPs. The EFG expected 50-100 entries by each LTP. **Table 1** lists the total number of uploaded metadata by the LTPs.

19 LTPs have provided the Country reports and inserted the expected number of metadata records until the end of 2016 and revised, completed the missing information of their EIGR records at the end of 2017. The information of the country reports summarized in this chapter is based on the 05/12/2017 status.

Table 1. Summary of the Inventory data collection and the total number of entries by EFG LTPs until the end of 2017 (taken from D2.3)

Country	Number of inserted metadata
1. Belgium/Luxembourg	62
2. Croatia	44
3. Czech Republic	584
4. Denmark	64
5. Finland	204
6. France	139
7. Germany	68
8. Greece	56
9. Hungary	54
10. Ireland	39
11. Italy	118
12. Netherlands	91
13. Poland	50
14. Portugal	55
15. Serbia	98
16. Slovenia	62
17. Spain	70
18. Ukraine	57
19. UK	68
Total:	1986

2.1 INFORMATION SOURCES

The LTPs used different sources on regional, national and international level to collect the relevant information for their EIGR entries. The LTPs were asked to classify their sources of information to the following groups.

- Institutions dealing with groundwater research/survey ,
- Groundwater monitoring, availability of data ,
- Journals/archives focused on hydrogeology,

The most important sources were the national databases, reports and journals, responsible governmental bodies, universities and national geological surveys. Table 2 contains the detailed summary of the sources of information per country as reported by the LTPs.

Table 2. Summary of the most important sources of information per country

Country	Source of information		
	Institutions dealing with groundwater research/survey	Groundwater monitoring, availability of data	Journals/archives focused on hydrogeology
Belgium/ Luxembourg	• DOV-Vlaanderen • Smart Geotherm • VITO • VITAQUA • University of Liege • University of Namur • University of Mons • ISSEP • SPAQUE • OVAM	• Hydrogeological Database of Wallonia • Geoportail • Service Geologique de Luxembourg • Region Wallonne DGO3-DGARNE • Automatic piezometric monitoring network	• www.belgium.iah.org
Croatia	• Croatian Geological Survey (HGI-CGS) • Faculty of Mining Geology and Oil Engineering (RGNF), University of Zagreb • Croatian waters (Hrvatske vode) d.d.	• Groundwater monitoring was conducted by a few experts from GHI-CGS, and also from Croatian waters • Local water management companies	
Czech Republic	• Section of Deputy Prime minister for Science, Research and Innovation • Ministry of the Environment of the Czech Republic • Ministry of the Industry and Trade of the Czech Republic • Czech Environmental Information Agency • Nature Conservation Agency of the Czech Republic • The Czech Science Foundation • TA CR • Czech Hydrometeorological Institute • Czech Geology Survey • T. G. Masaryk Water Research Institute	• Czech Environmental Information Agency • Nature Conservation Agency of the Czech Republic • Czech Geology Survey • T. G. Masaryk Water Research Institute • Czech Hydrometeorological Institute	• National repository of grey literature • Geofond • Geopub • Information Register of R&D result • The Central Register of R&D projects • National Library of Technology
Denmark	• Geological Survey of Denmark		
Finland	• Geological Survey of Finland (GTK) • University of Helsinki, Finland • Finnish Environment Institute (SYKE) • Regional Centres for Economic Development, Transport and the Environment • Radiation and Nuclear Safety Authority • Regional Water Supply Enterprises • Ramboll Finland Oy (Ltd.)	• Geological Survey of Finland (GTK) • Finnish Environment Institute (SYKE) • Regional Centres for Economic Development, Transport and the Environment	• HAKKU (archives of Geological Survey of Finland) • HELDA (archives of Helsinki University) • OIVA (archives of the Finnish Environment Institute)

France	- French Geological Survey - www.cordis.europa.eu	ADES database for groundwater	- Science Direct - Springer - Google Scholar
Germany	- www.umweltbundesamt.de - www.bgr.bund.de	www.bgr.bund.de	Grundwasser (Springer journal)
Greece	Institute of Geology and Mineral Exploration	- Aristotle University of Thessaloniki - National Documentation Centre	Special Secretariat of Water
Hungary	- Geological and Geophysical Institute of Hungary (MAFI) - National Archive of the Office of Mining and Geology	Geological and Geophysical Institute of Hungary (MAFI)	- Acta Geologica Hungarica - National Archive of the Office of Mining and Geology
Italy	- ISPRA (National Institute for Environmental Protection and Research) - ISTAT (National Statistics Institute) - IRSA-CNR (Water Research Institute of the National Research Council) - Regional and Basin Authorities	Interregional and regional monitoring networks	- Italian Journal of Groundwater - Italian Journal of Engineering Geology and Environment - L'Acqua journal
Netherlands	- Geological Survey of the Netherlands - Deltares - Alterra - KWR Watercycle Research - Netherlands Hydrological Instrument - Utrecht University - VU University Amsterdam - Wageningen University and Research - Dutch Provinces - Dutch Water Boards	DINOloket	- Stromingen Journal - NJG: Netherlands Journal of Geosciences - H2O Journal, by KNW
Poland	- Polish Hydrogeological Survey - Ministry of Environment - Polish Geological Institute - Chief Inspectorate of Environmental Protection - Universities		
Portugal	- Universidade dos Açores - Universidade do Algarve - Universidade de Lisboa - Universidade Lusíada - Universidade NOVA de Lisboa - Universidade da Madeira - Universidade do Porto	- APA: Portuguese Environmental Agency - SNIAMB	International Journals (Journal of Hydrology, Chemie der Erde, Engineering Geology, Geofísica Internacional, Journal of Volcanology and Geothermal Research, Agricultural Water Management; Chemosphere, Science of the Total Environment, Estuarine, Coastal and Shelf Science, Geothermics, Journal of Geochemical

	• Universidade de Trás-os-Montes e Alto Douro • APRH - Associação Portuguesa dos Recursos Hídricos • Universidade de Aveiro • Universidade de Coimbra • Universidade da Covilhã • Instituto Politécnico de Beja • Instituto Politécnico de Castelo Branco • Instituto Politécnico do Porto • Câmara Municipal do Porto • CCDR-Algarve • DGEG – Direção Geral de Energia e Geologia • Direção Regional do Ambiente - Açores • EMAS Beja • LNEG – Laboratório Nacional de Energia e Geologia • SMAS Ponta Delgada		Exploration, Applied Geochemistry, Sensors and Actuators A: Physical • Project Report (WAT – Water and Territories) • International Proceedings Journals (Procedia Engineering, Procedia Earth Science and Planetary Science, Materials Today: Proceedings) • National Journals (Recursos Hídricos (APRH), GEONOVAS (APG))
Serbia			• National journals (Vodoprivreda, Tehnika, Anali Balkanskog poluostrva) • International journals (Hydrogeology Journal, Environmental and Earth Science, Archives of Mining Sciences) • Papers presented at national and international conferences in the fields of geology and hydrogeology.
Slovenia	• Slovenian Geological Survey • Ministry of Environment		
Spain	• Geological Survey of Spain (IGME)	• Public information from monitoring networks	• Research Gate platform
Ukraine	• EA UAG • Institute of Geological Sciences of NAS of Ukraine • Institute of Geology • Taras Shevchenko National University • Oles Honchar Dnipropetrovsk National University • State Geological Survey of Ukraine • "State Informational Geological Fund of Ukraine" • V.N. Karasin Kharkiv National University • M.P.Semenenka Institute of Geochemistry, Mineralogy and Ore Formation of NAS of Ukraine		• Geological Journal • Dnipropetrovsk University bulletin. Geology, geography • Bulletin of Taras Shevchenko National University, Series "Geology" • V.N. Karasin Kharkiv National University Bulletin, series "Geology, Geography, Ecology" • Geochemistry and Ore Formation Journal • Maksymovych Scientific Library • Vernadsky National Library of Ukraine

2.2 TYPE OF INFORMATION

The LTPs classified the information they gathered according to the previously defined research and knowledge classes (**Figure 1**). Most of the data from universities and research institutes were ranked as Class 1 or Class 2 and the information gathered from the regional authorities usually valued as Class 3 and Class 4.

The majority of the metadata (more than 900 entries) are related to Class 4, mainly because the Czech LTP inserted information on a large number of projects (583 entries) to the EIGR which were all ranked as Class 4. The number of peer-reviewed articles in scientific journals (Class 1) and the number of reports from research projects and publications in national technical journals (Class 3) has a similar occurrence (~24%). The Class 2 resources have the lowest number of EIGR records inserted by LTPs, less than 10 % were classified into Class 2. The reason for that, is that the LTPs focus on publications and data sources which are not already available through the well-known and most appreciated research databases (i.e. Web of Science and Scopus).

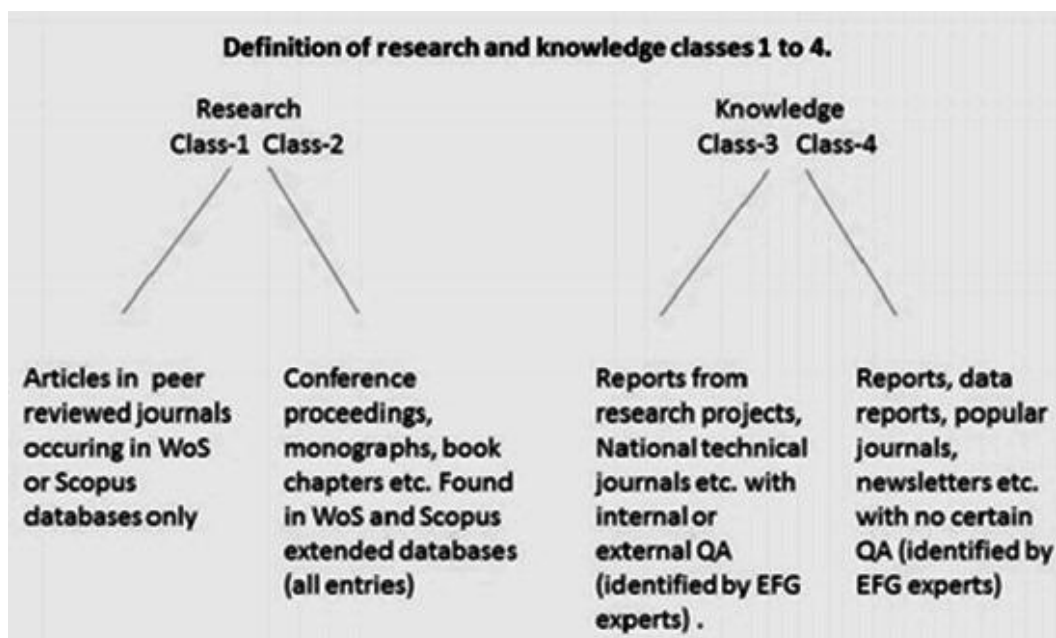


Figure 1. Research and knowledge classes identified in the KINDRA project

2.3 TYPES OF DATA AND RESOURCES (TOPICS OF THE DATA)

The LTPs grouped the resources, they inserted to EIGR into the following 7 data types or resource categories. The type of the resources included in this wide range of information with different accessibility and formats are:

- a) National databases;

- b) National and local reports containing facts and data;
- c) Hydrogeological maps;
- d) Technical reports, guidelines, manuals, etc.;
- e) Books and book chapters;
- f) Position papers and/or important papers on PR journals;
- g) Others.

The number of entries per LTP related to each of the data types or resource categories were summarized in **Table 3** (taken from D2.3). These numbers were reported in the Country reports.

Figure 2 shows the distribution of the type of resources. The most dominant resource category is the “National and local reports containing facts and data” covers almost the half (45,3%) of the metadata. The “Position papers and/or important papers in peer reviewed journals” data type has also a significant number of entries 29,2%. The remaining 25,5% distributed between the “National databases” (3,1%), “Hydrogeological maps” (4,1%), “Technical reports, guidelines, manuals, etc.” (3,3 %), “Books and book chapters (6,1 %)” and “Other” (9%) topics.

Table 3. Summary of the topics of data with the numbers of entries reported by the LTPs
(source: D2.3)

Country	DATA TYPE/RESOURCE CATEGORY						
	Databases	Reports	Maps	Technical reports	Books	Papers	Others
Belgium/ Luxembourg	3	9	4	2	2	7	2
Croatia	-	-	1	1	1	2	-
Czech Republic	12	587	11	10	55	-	98
Denmark	3	8	4	-	-	5	12
Finland	3	78	4	15	10	2	-
France	2	2	2		1	82	8
Germany	-	4	3	-	-	9	-
Greece	1	10	-	-	-	-	3
Hungary	2	8	3	-	-	-	-
Ireland	-	6	-	6	-	27	-
Italy	5		3	1	-	87	3
Netherlands	4	20	21				7
Poland	-	16	6	9	17	-	2
Portugal	-	-	-	1	-	48	9
Slovenia	1	3	-	6	10	42	-
Serbia	1	7	-	-	3	87	-
Spain	4	4	1	3	5	28	4
Ukraine	9	12	6	-	-	24	-
UK	2	-	1	2	-	49	5

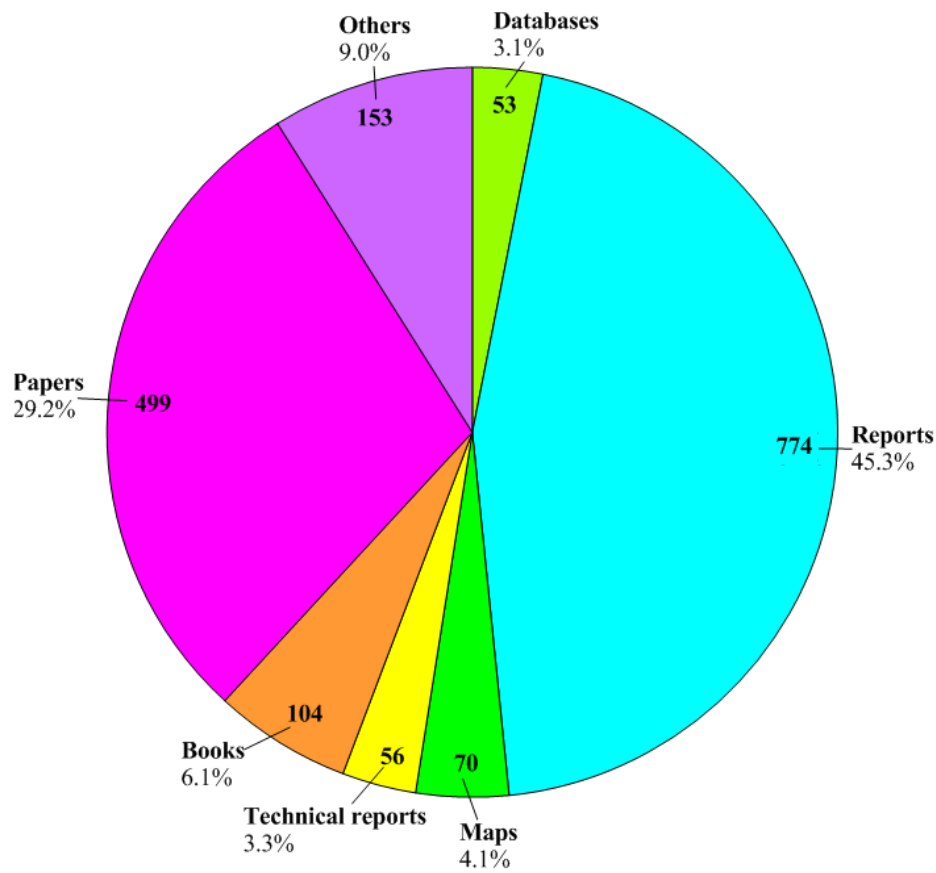


Figure 2. The distribution of data types/resource categories in the EIGR entries reported in the country reports

2.4 GAPS AND BARRIERS

The barriers to find the data and the gaps in finding the information are different in each country. The barriers reported the most by LTPs are (1) the language as most of the relevant information is in national languages and (2) the confidential-copyright issues. Another barrier is the scattering of data among national, regional and local authorities, as there is typically no national public database or the data concerning the groundwater is outdated.

In case of the Czech LTP, who provided the most of the entries to EIGR, the main barriers are (1) the classification and selection of relevant data due to its abundance and (2) the language issues, resulting in that Czech is the dominant language of the reported data.

Many of the LTPs haven't reported gaps (Finland, Denmark, Germany, Hungary, and Ukraine). In case of the other LTPs the main gaps in finding the data are the scattering of information-which makes the data gathering more complex and time consuming as a huge amount of information exists at various sources that are hard to find and review the past and recent hydrogeological research on national level.

The Czech LTP identified the abundance of relevant data and limited human resource as a gap.

The detailed description of gaps and barriers for each country are described in D2.3 (Country reports).

2.5 OVERVIEW OF EIGR METADATA

The number of EIGR records shows a significant increasing trend from 15/6/2016. Until 28/02/2018 2330 metadata were inserted to the EIGR by LTPs and the KINDRA consortium partners. From the records, 2178 classified into the HRC-SYS three main categories, the data assessment based on these records. The development of EIGR is illustrated in **Figure 3**.

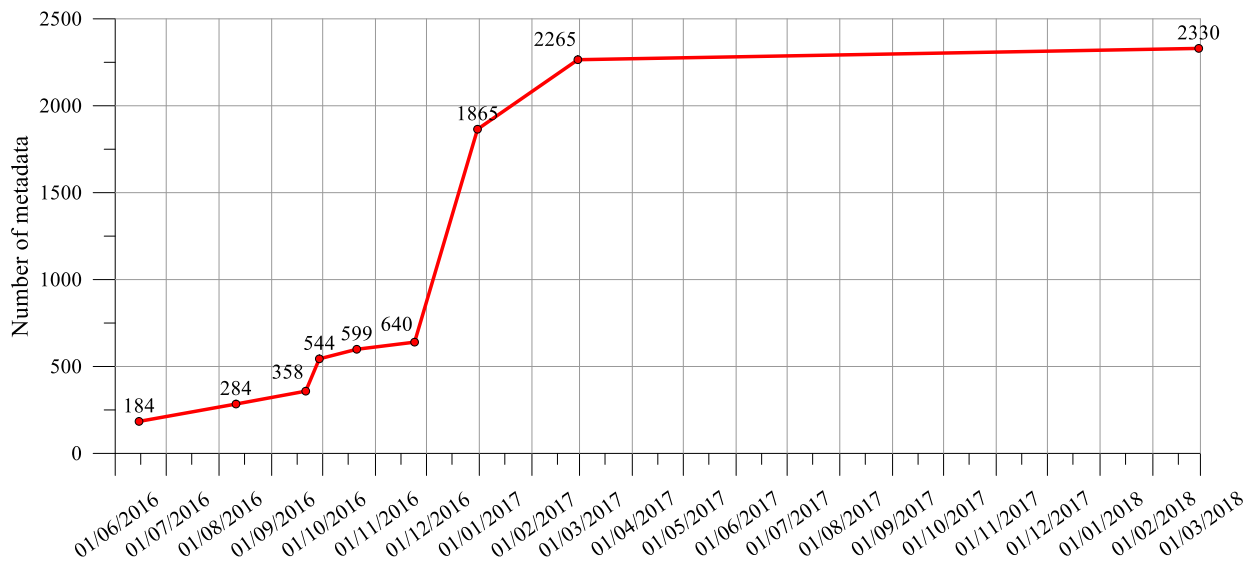


Figure 3. The Total number of EIGR contents inserted by the LTPs and KINDRA consortium partners until 28/02/2018

3 ANALYTICAL TOOLS FOR DATA ASSESSMENT

3.1 DEVELOPED EIGR TOOLS FOR DATA ANALYSIS AND VISUALISATION

This chapter describes the three EIGR tools developed for analyzing and visualizing groundwater related research and information stored in the EIGR database. These tools are:

- “Keywords cloud”
- “Resource Distribution Map”
- “Topics 2D Chart”

The **Keywords Cloud** section indicates the relative importance of keywords, it shows the 10 most popular keywords in the EIGR, related to the records which have been uploaded to the Inventory. Clicking on each one of these keywords will take users to the selection of records that include them (**Figure 4**).

The dominant keyword in EIGR metadata is the “Climate, environment and resources” with 1903 hits, closely followed by “Geology” with 1655 records. The size and colors of the fonts are related to the popularity of the keywords, the frequent keywords have dark red colour and larger fonts, whereas the keywords with lower number of hits have light orange colour and small font size.

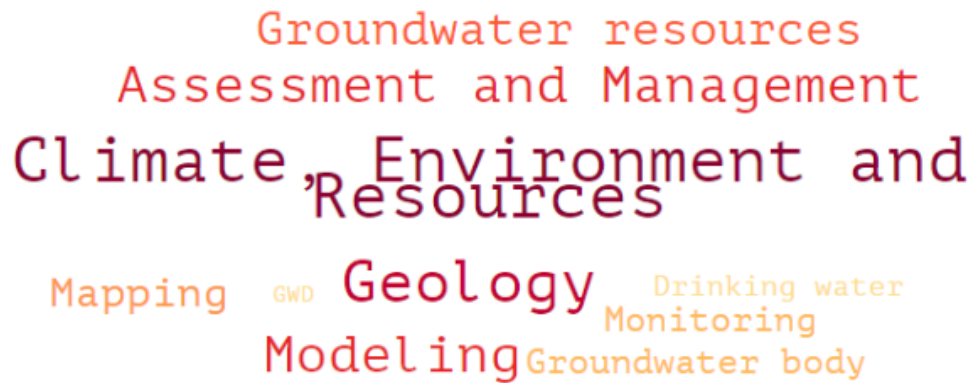


Figure 4. The Keywords Cloud

The **Resource Distribution Map** will show the number of contributions to the Inventory by country. When selecting a specific territory this tool shows all the resources that were uploaded to the EIGR which have been generated by a specific country (**Figure 5**).

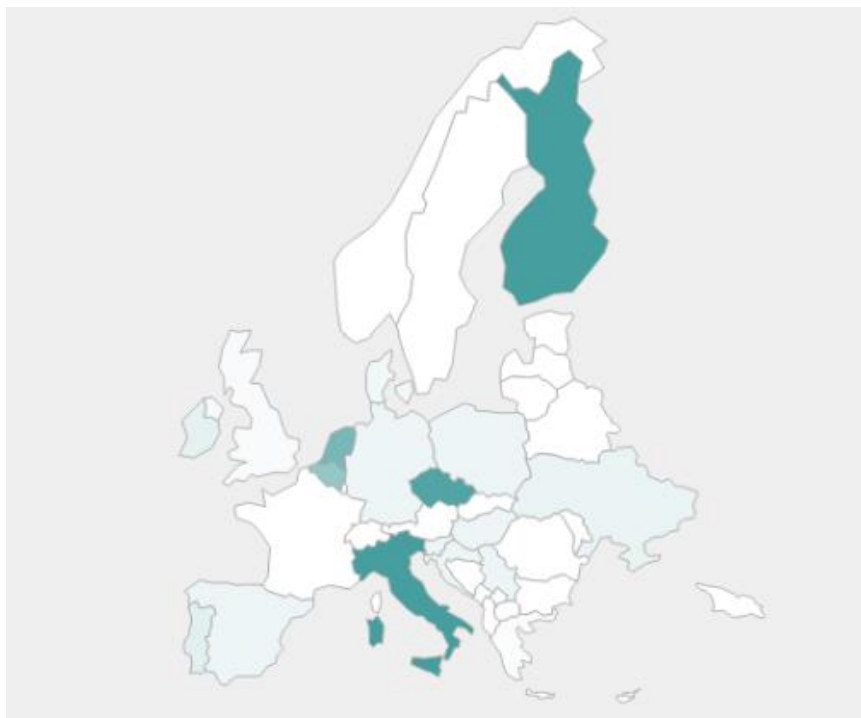


Figure 5. The Resource Distribution Map

The section related to the Research Topics (RT)-Operational Actions (OA) 2D Chart is a quick tool, which allows all users to quickly view the number of resources classified according to the main Hydrological Research Classification System (HRC-SYS) and the three overarching categories: Operational Actions, Research Topics and Societal Challenges. When selecting a specific Societal Challenge from the drop-down menu, the 2D Chart will represent the number of resources which has been uploaded to EIGR and classified for the corresponding Research Topic (RT) - Operational Actions (OA) Categories. Users will be able to access the specific records that match the criteria by clicking on the values represented in the graph (**Figure 6**).

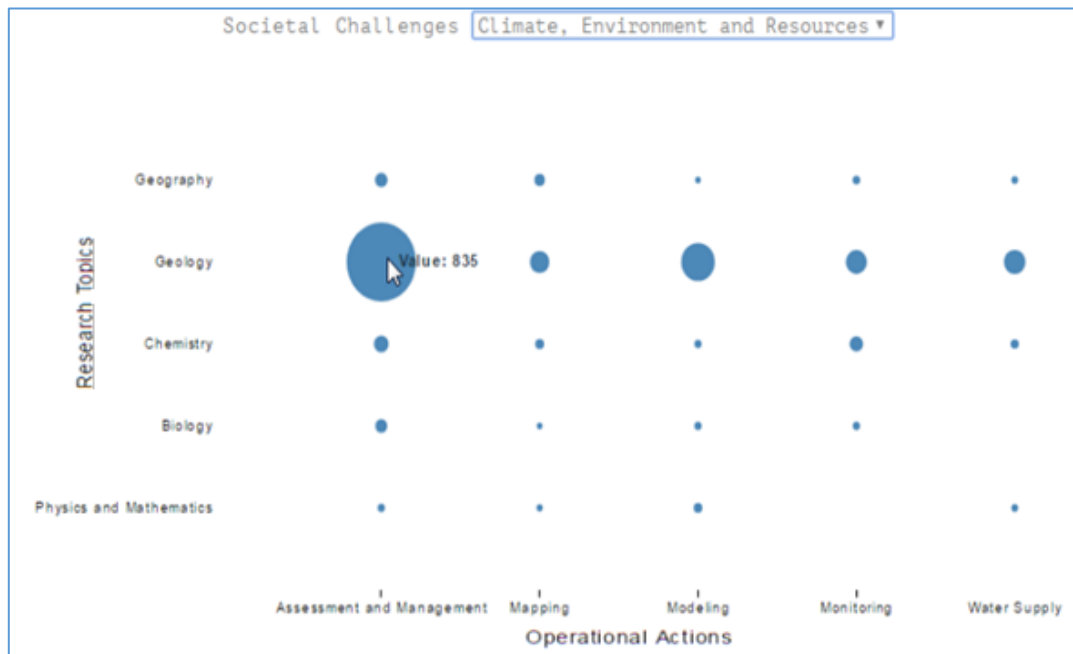


Figure 6. The Research Topics -Operational Actions 2D Chart (for each of the five Societal Challenges) of EIGR

3.2 EXTERNAL TOOLS

External, standalone tools can also be used for EIGR data analysis, but the metadata must be exported in an appropriate format to import data to such tools. From the different kind of available external tools, this chapter describes the bibliographic map features of the freely available, open source visualization tool, the VOSviewer. The VOSviewer (<http://www.vosviewer.com>) was created for bibliometric networks, developed at Leiden University, The Netherlands. The software requires Java version 6 or higher to be installed on the system. The user's manual for this tool is also available on the www.vosviewer.com website.

The **Heat map** feature gives a quick overview of the co-occurrence of the most often used keywords. In this view the keywords are represented by their labels. Each of these keywords in a map has a colour, ranging from blue to red, which depends on the density of keywords at that point. The larger the number of keywords in the neighborhood of a point and the higher

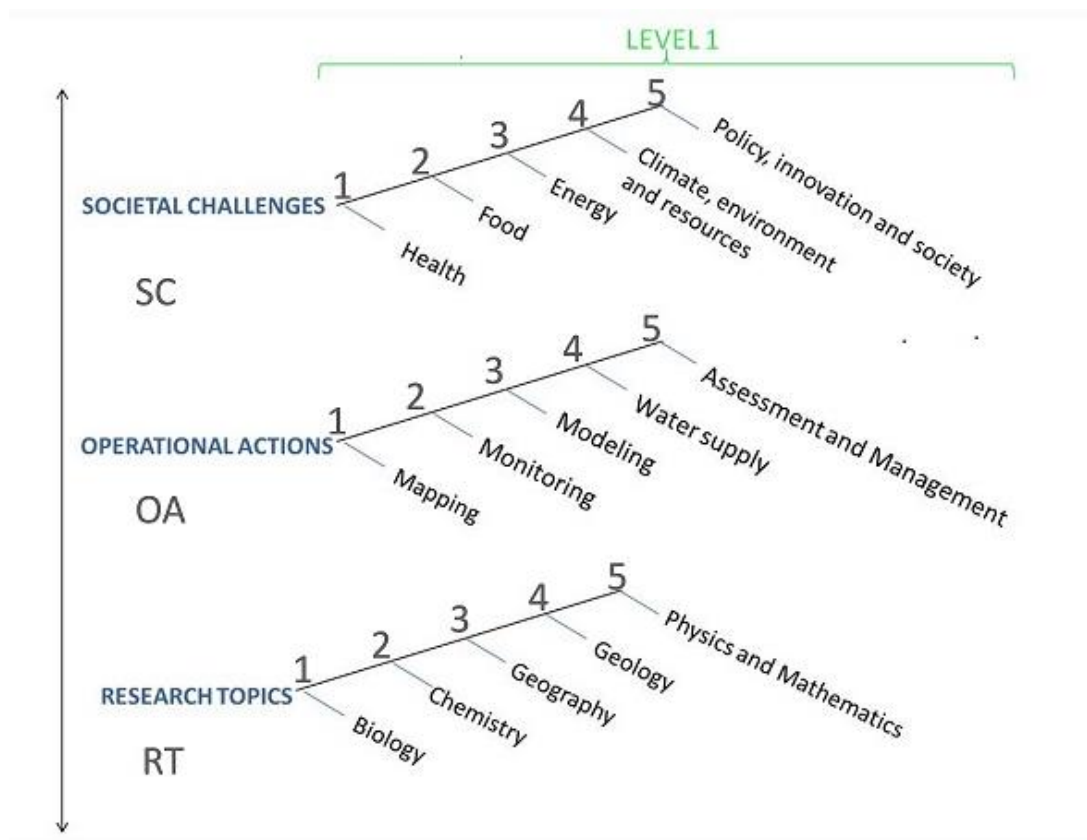


Figure 9. The HRC-SYS three overarching categories: 1) Societal Challenges, 2) Operational Actions and 3) Research Topics and their five main sub-categories (“keywords”) used for classification of the resources uploaded to EIGR.

The majority (1903 entries, 87%) of the metadata uploaded to EIGR were related to and classified in **Societal Challenges** (SC), Climate, environment and resources. The number of records in the other four challenges (i Health, ii Food, iii Energy and iv Policy, innovation and society) varies between 1%-7% (**Figure 10**).

In total 2178 metadata were classified into **Operational actions** (OA) sub-categories as illustrated in **Figure 11**. More than the half (53%) of EIGR metadata (1165 entries) were classified into the Assessment and management category, and the Modeling actions also has a relative high number of entries (397 entries, 18%). The OAs “Mapping” (164 entries) and “Water supply” (215 entries) categories have a similar number of entries classified in HRC-SYS. For the **Research topics** (RT), “Geology” (1655 entries) is the dominant category, with the 76% of the total number of records (2178 entries). The “Biology”, “Geography” and “Physics and Mathematics” collectively reached 266 entries, 12 % of the total number of metadata (**Figure 12**).

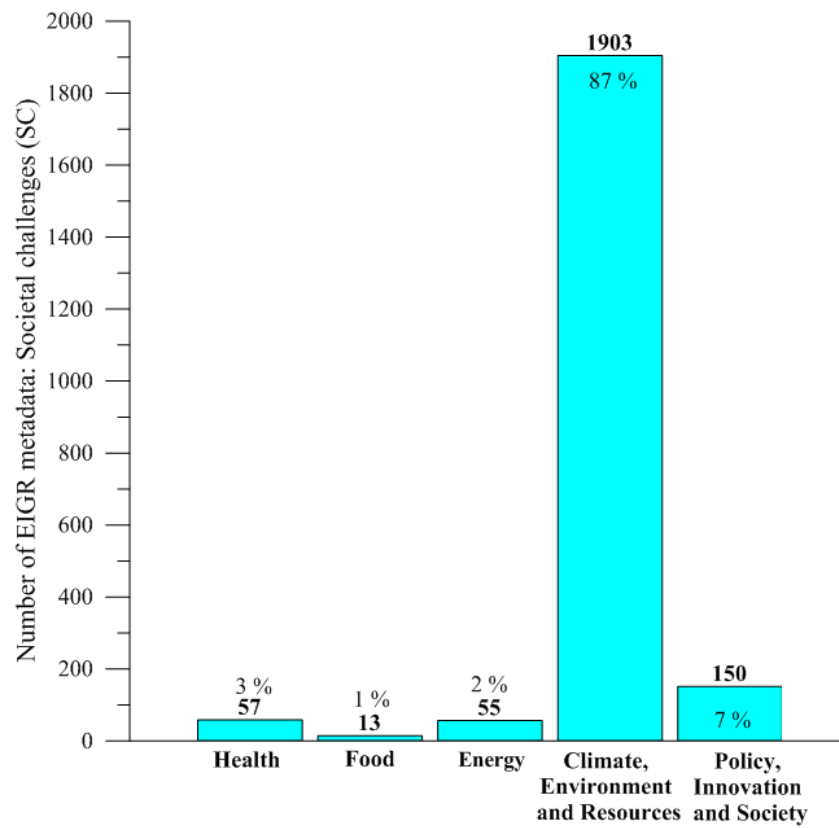


Figure 10. The distribution of EIGR metadata related to Societal Challenges (SC)

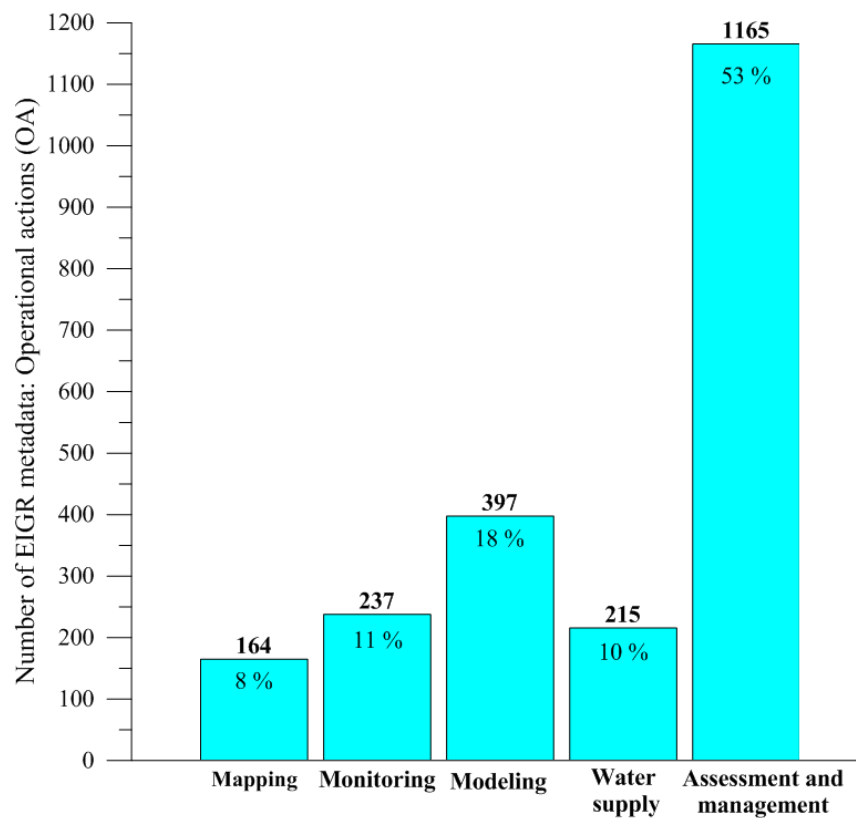


Figure 11. The distribution of EIGR metadata related to Operational Actions (OA)

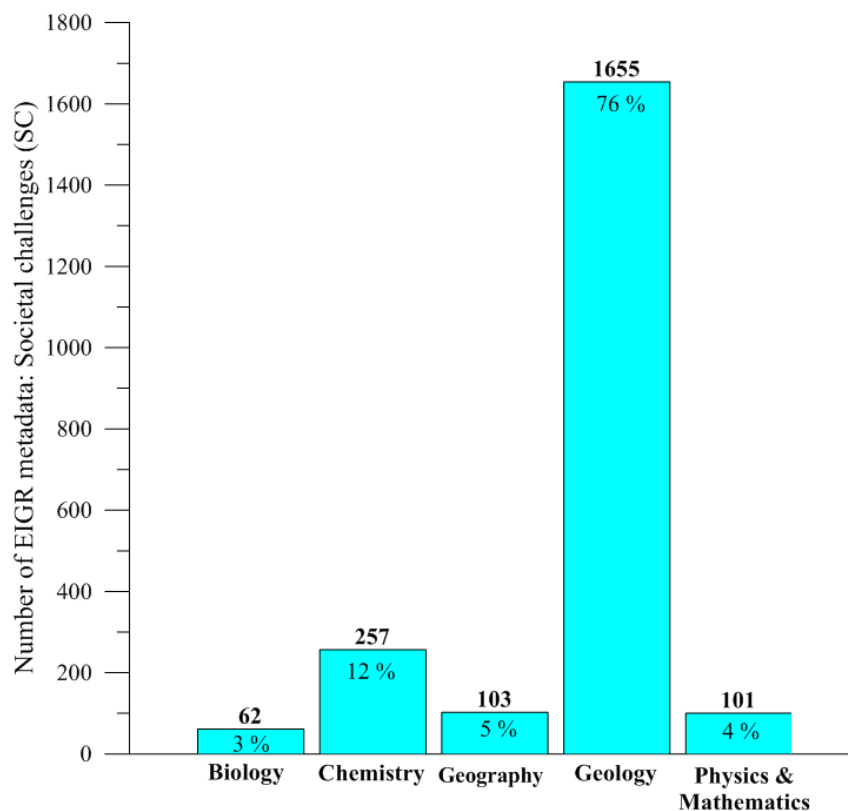


Figure 12. The distribution of EIGR metadata related to Research Topics (RT)

4.1.1 Co-occurrence analysis of HRC-SYS categories by EIGR metadata

As previously described, each of the 2178 metadata were classified into 3 of the HRC-SYS categories to which they primarily relate i.e. one societal challenge (SC), one operational action (OA), and one research topic (RT) by the national experts (LPTs) uploading metadata and information on the resources, allowing a co-occurrence analysis of these categories and estimation of the number of studies within the combination of these categories based on the uploaded metadata.

The following graphs are made separately outside EIGR and based on the 28/02/2018 existing EIGR data and the EIGR **2D Chart (Figure 6)** tool, which developed to perform the co-occurrence analysis of EIGR records.

Figure 13-17 shows the number of records which comply with the three specified HRC-SYS categories. The size of the bubbles relates to the number of the metadata, which belongs to the three defined categories. The gaps, i.e. combinations of categories with no studies in EIGR are indicated as 0 value.

For Societal challenges (SC) categories the number of metadata belonging to Research topics (RT) and Operational actions (OA) are plotted.

In **Figure 13** the Research topics (RT) and Operational actions (OA) are plotted for a *Societal challenges (SC)-Health* category. 57 EIGR records were assigned to the *Health* challenge and

from this value, the 21% of the records belong to *Operational Actions-Water supply, Research Topics-Chemistry* and *Societal challenges-Health* categories. In this intersection 28 Kindra Thesaurus keywords can be found, as specified below. 10 gaps can be found on the chart.

Aquifer	Development country	Geochemistry
Aquifer vulnerability	Water supply	Geology
Contamination	Arsenic	Groundwater resources
Drinkable water	Assessment	Radon
Geology	Environment	Industry
Groundwater body	Health	Human toxicology
Human health	Lead	Legislation
Laboratory measurements	Metals	Technique
Quality	Removal	
Chemistry	Africa	

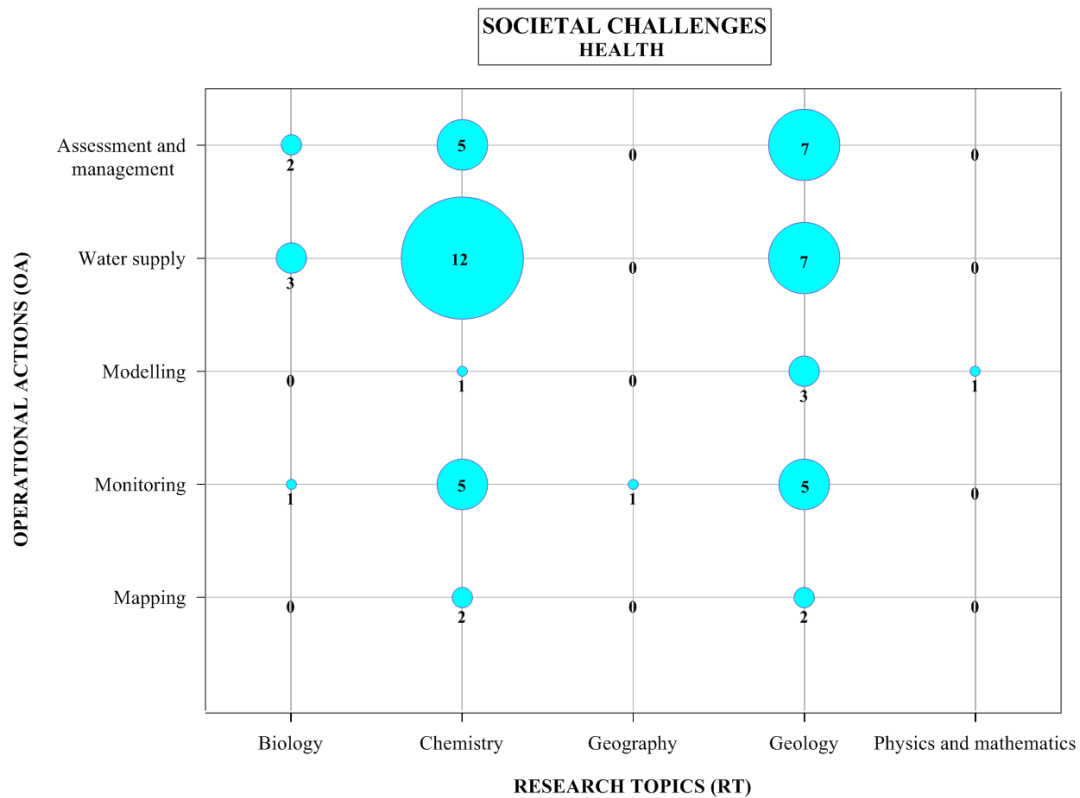


Figure 13. The number of metadata for Societal challenges -Health and 5-5 classes of Research topics and Operational actions

Figure 14 shows the intersections between the Research topics (RT) and Operational actions (OA) categories for the *Societal Challenges-Food* category. Only 13 records belong to the *Societal Challenges-Food* action. Most of the records (30%) were assigned to *Operational Actions-Monitoring* and *Research Topics-Geology* and *Societal challenges-Food* categories. In this intersection 2 keywords can be found, which are: Food, and food production. In 17 intersections gaps can be found.

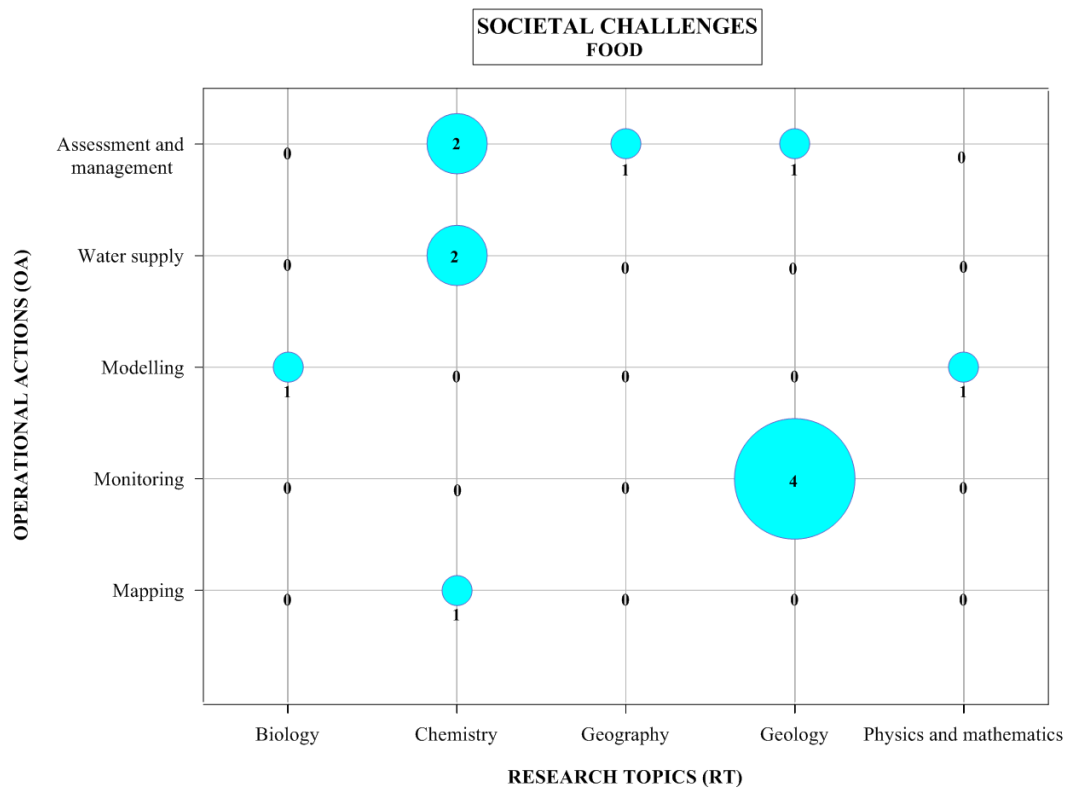


Figure 14. The number of metadata for Societal challenges -Food and 5-5 classes of Research topics and Operational actions

Figure 15 illustrates the Operational actions (OA) and Societal challenges (SC) for the *Societal Challenges (SC)-Energy* category. 55 EIGR records were assigned to the *Energy* challenge. In the figure, it could be seen, that the majority of the entries were classified into the *Operational Actions-Assessment and management* and *Research Topics-Geology* categories. In case of 11 intersections, there are research gaps. In active intersections, 17 Kindra Thesaurus keywords can be found:

hazard	conceptual model	groundwater resources
urban areas	Europe	borehole logging
geothermal energy	energy	characterisation
groundwater body	energy production	urban groundwater
shale gas	hydrochemical modeling	salinization
coastal aquifer	pump test	

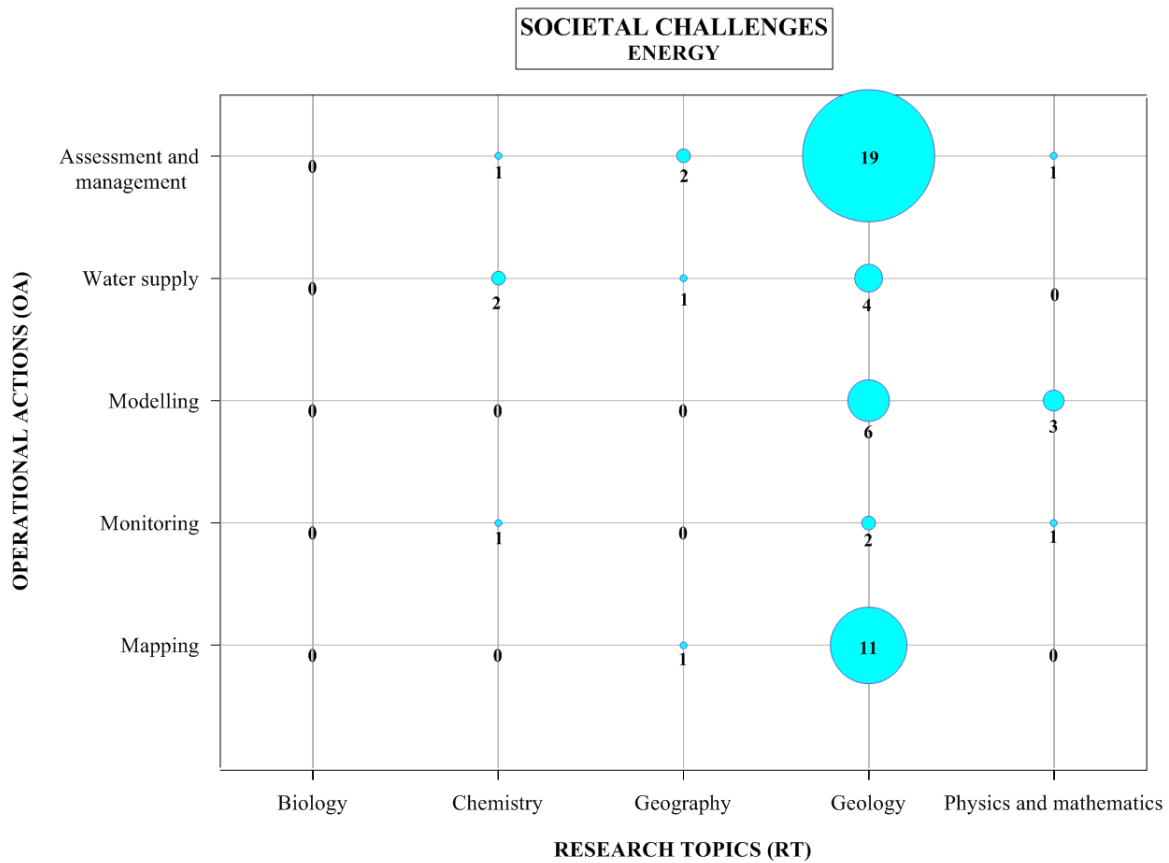


Figure 15. The number of metadata in the intersections of Societal challenges -Energy and 5-5 classes of Research topics and Operational actions

Figure 16 shows the intersections between the Research topics (RT) and Operational actions (OA) categories for the *Societal challenges- Climate, environment and resources* category. 1903 records were assigned the *Societal challenges- Climate, environment and resources* challenge. The dominant intersection with 889 records (47%) is the *Operational Actions-Assessment and management* and the *Research Topics-Geology*. In this intersection more than 1000 Kindra Thesaurus keywords can be found, the 10 most frequently occurred are:

groundwater resources	modeling	aquifer
groundwater body	GWD	quality
monitoring	hydrology	drinking water
management		

There is not research gaps in this cross section plot, demonstrating the activity on this topics by hydrogeological research and knowledge.

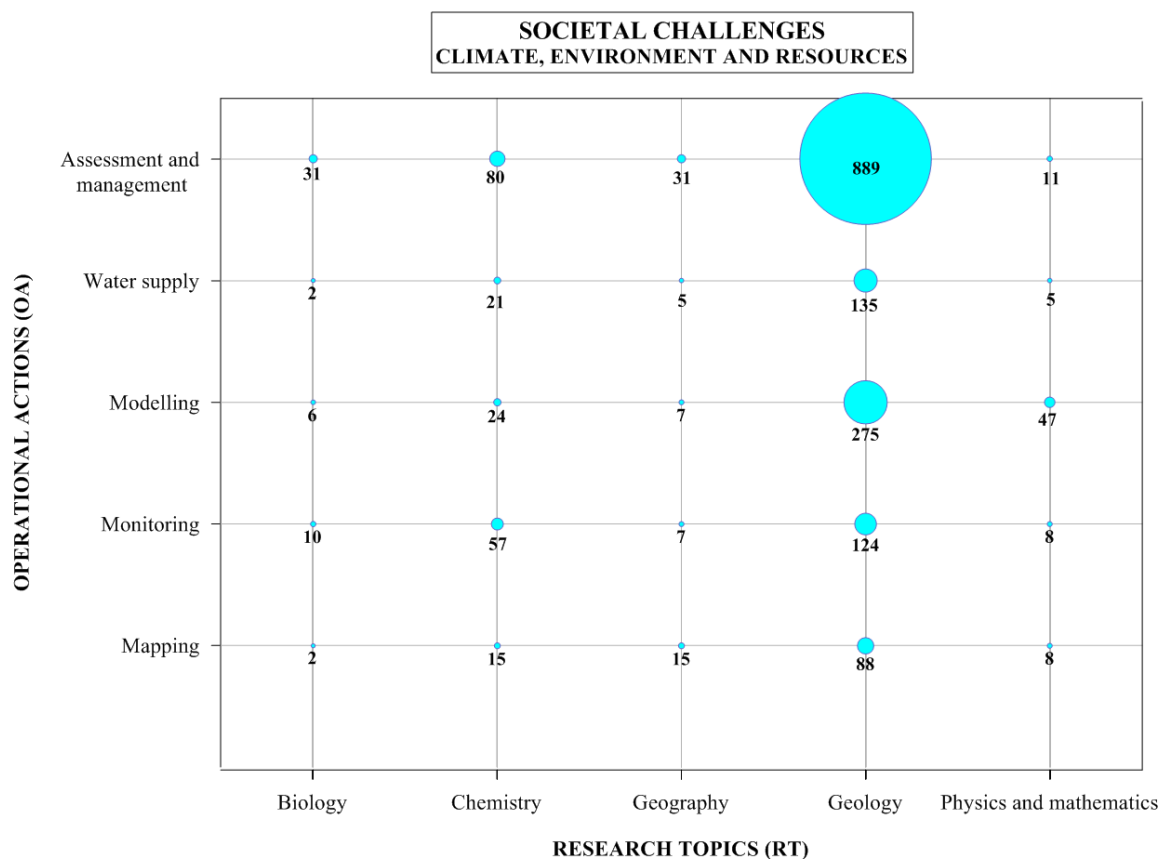


Figure 16. The number of metadata in the intersections of Societal challenges -Climate, environment and resources and 5-5 classes of Research topics and Operational actions

The last figure from this co-occurrence chart series is **Figure 17**, it shows the intersections between the Research topics (RT) and Operational actions (OA) categories for the *Societal Challenges (SC)-Policy, innovation and society* category. 150 records are assigned to this challenge. The dominant intersection with 30% of the records is in the *Operational Actions-Assessment and management* and the *Research Topics-Geology* categories. In this intersection more than 100 Kindra Thesaurus keywords can be found, the 10 most frequently occurred are:

abstraction	legislation	groundwater body
GWD	management	integrated management
hydrology	mapping	policy
groundwater resources		

From the 25 possible intersections of HRC-SYS categories, 5 intersections indicate research gaps for which no records have been classified.

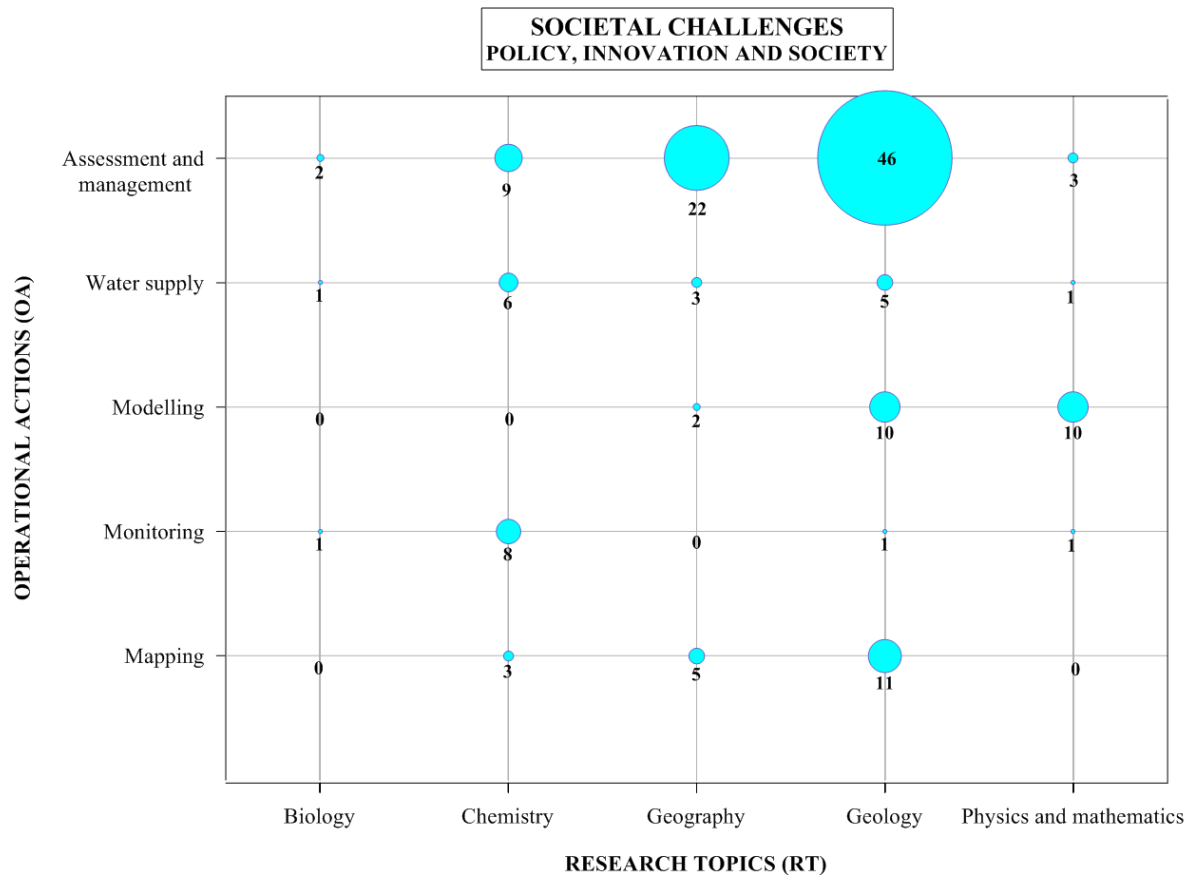


Figure 17. The number of metadata in the intersections of Operational actions-Water supply and the 5-5 classes of Research topics and Societal challenges

As the 2D chart series (**Figures 13-17**) illustrated, the majority (889) of the EIGR records were assigned to the *Operational Actions-Assessment and management* and *Research Topics-Geology* and *Societal challenges-Climate, environment and resources* categories. In case of the other Operational actions (Mapping, Monitoring, Modeling, Water supply), also the *Research Topics-Geology* and *Societal challenges-Climate, environment and resources* co-occurrences are dominant. In many intersections (43 from the possible 125 intersections) research gaps can be found based on the EIGR metadata.

The detailed investigation of the co-occurrence of categories can be found in D3.3 (Gaps and trends in groundwater research) deliverable.

4.1.2 The keyword occurrence by country

The 10 most frequent keywords occurring in the records uploaded to EIGR were different in every country.

In **Figure 18**, “keyword clouds” can be found for the most commonly occurred keywords in the entries uploaded by Belgium, Denmark, France and Finland. The information was derived from the EIGR database.

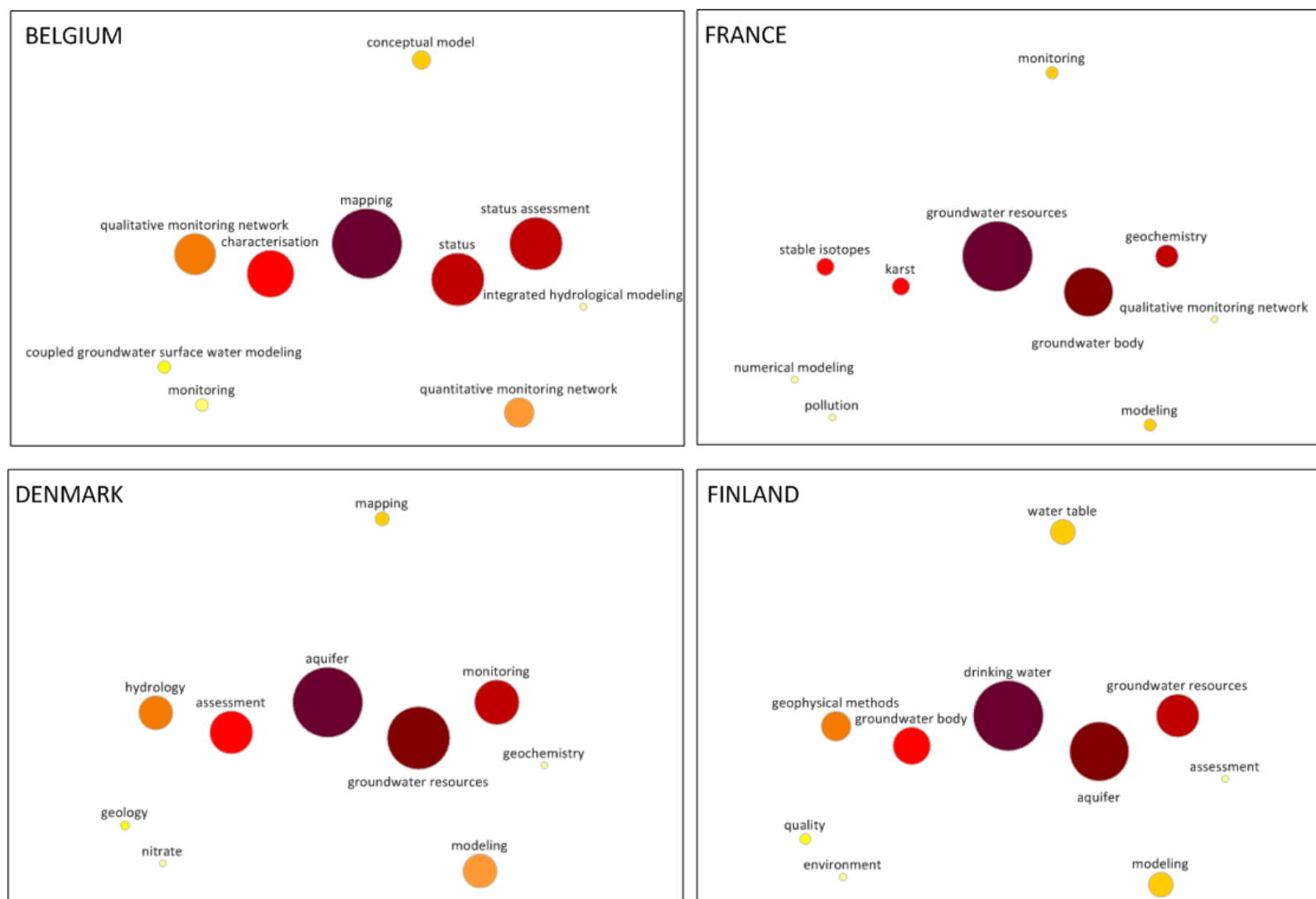


Figure 18. The 10 most popular keywords for four countries

The size of the bubbles related to the number of occurrences of each keyword and has a relative scale of the diameter. The color of the bubbles also related to the number of occurrences as in case of the KINDRA EIGR “Keywords cloud” tool.

The most frequent keywords (mapping, groundwater resources, aquifer, drinking water) in the uploaded entries are different in every country, but there are lot of keywords, which are occurred more times (e.g modeling, aquifer, groundwater resources). The most of the uploaded entries are related to reports of national -or international-research projects, and the research topics are site specific, that could be the reason of the differences in the most frequent keywords country by country.

The **Figure 19** summarized the keywords occurrences for 20 countries, which were uploaded resources to EIGR. It can be seen, that the *groundwater resources* and the *monitoring* keywords were appeared in the EIGR entries of 10 nations, these keywords have a relative great importance. The distribution for the other keywords are much more uniform.

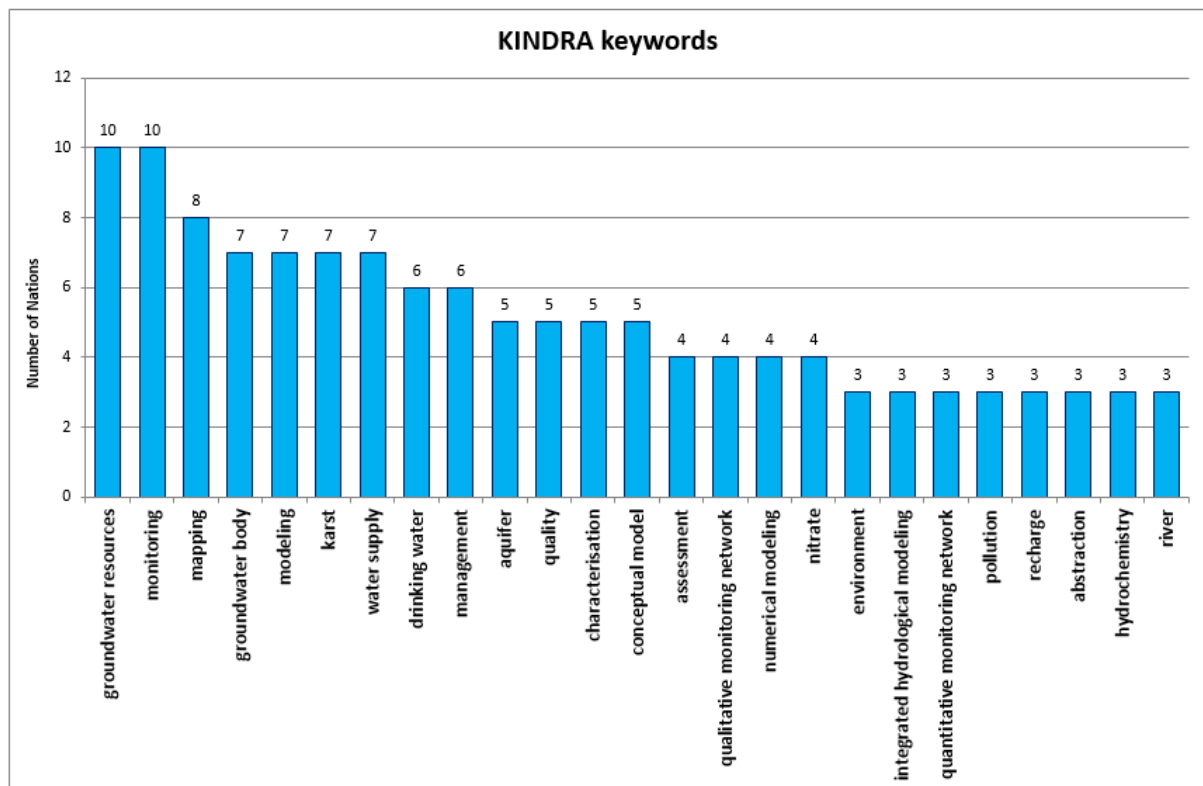


Figure 19: The keyword occurrences vs. the number of nations

The keyword cloud on **Figure 20** shows the 10 most popular keywords in the EIGR. It can be seen that the *groundwater resources* keyword is dominant, and it can be found in 412 uploaded documents can be found. The occurrence of the 10 most popular keywords is summarized in **Table 4**.

The table also contains the most four frequent co-occurred keywords for each “popular keyword”. In parentheses the number of co-occurrence can be seen in EIGR documents.

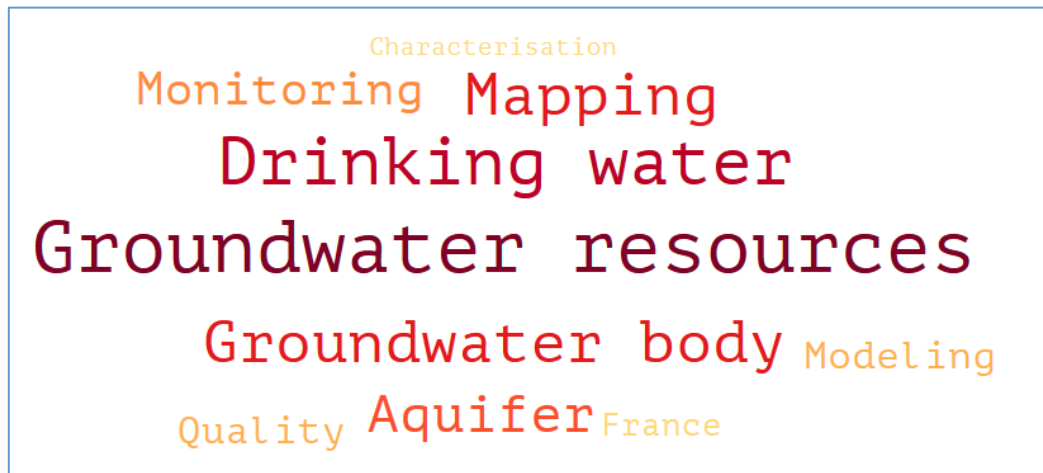


Figure 20: The most popular keywords in the EIGR (on the status 28/02/2018)

Table 4. The 10 most popular keywords in the EIGR entries and their four frequent co-occurrence keywords

Keywords	Nr. of occurrences in EIGR entries	Co-occurrences with other keywords
Groundwater resources	412	Groundwater body (113) Aquifer (99) Drinking water (96) Modeling (76)
Drinking water	220	Groundwater resources (96) Aquifer (84) Groundwater body (76) Geophysical methods (64)
Groundwater body	198	Groundwater resources (113) Aquifer (80) Drinking water (76) Geophysical methods (68)
Mapping	154	Groundwater resources (54) Drinking water (40) Aquifer (31) Qualitative monitoring network (29)
Aquifer	174	Groundwater resources (99) Drinking water (84) Groundwater body (80) Geophysical methods (70)
Monitoring	171	Qualitative monitoring network (41) Quality (40) Groundwater resources (37) Quantitative monitoring network (36)
Modeling	157	Groundwater resources (76) Aquifer (70) Groundwater body (68) Geophysical methods (65)
Quality	154	Drinking water (55) Groundwater resources (51) Monitoring (40) Characterisation (29)
Characterisation	96	Quality (29) Monitoring (28) Conceptual model (27) Mapping (27)
France	95	Groundwater resources (18) Geochemistry (15) Stable isotopes (14) Groundwater body (14)

4.2 DATA ASSESSMENT FOR RESEARCH AND KNOWLEDGE CLASSES

Of 2178 published EIGR records all of the entries were classified into the defined research and knowledge classes (**Figure 1**). **Figure 21** shows the distribution of metadata between Class 1- Class 4.

43,3 % of the metadata (943 entries) are related to Class 4. The number of peer-reviewed articles - ranked as Class 1 - and the number of reports from research projects, national technical journals (valued as Class 3) have similar occurrences, 23,5% and 24,3%.

Class 2 has the lowest number of EIGR records (193 entries, 8,9%) as in case when only the LTP's metadata were taken into consideration (Chapter 2.2). It should be noted that the dataset is biased and the main reason for the observed metadata distribution is that, the users (LPTs) were asked to focus on projects, maps, reports and data sources which are not available in existing well-known and accessible scientific databases (e.g. Web of Science or Scopus).

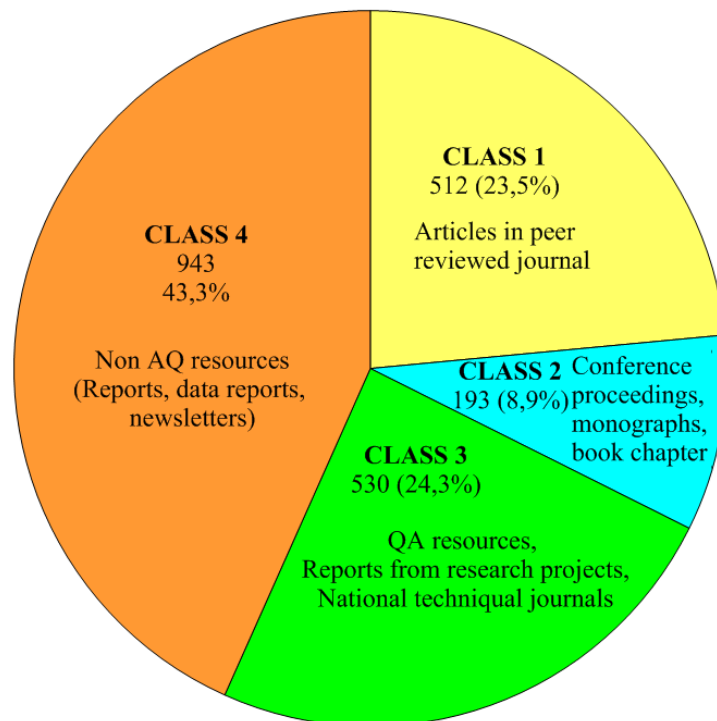


Figure 21. The distribution of EIGR metadata related to Research and knowledge classes (KC)

4.3 DATA ASSESSMENT FOR TECHNOLOGICAL READINESS LEVEL CLASSES

The EIGR database contains 2178 published records, from which 1590 were assigned to one of the nine Technological Readiness Levels (TRL):

- TRL 1: Basic principles observed;
- TRL 2: Technology concept formulated;
- TRL 3: Experimental proof of concept;
- TRL 4: Technology validated in lab;
- TRL 5: Technology validated in relevant environment (industrially relevant environment in the case of key enabling technologies);
- TRL 6: Technology demonstrated in relevant environment (industrially relevant environment in the case of key enabling technologies);
- TRL 7: System prototype demonstration in operational environment;
- TRL 8: System complete and qualified;
- TRL 9: Actual system proven in operational environment (competitive manufacturing in the case of key enabling technologies; or in space).

The **Figure 22** illustrates the distribution of metadata between the TRLs.

The majority of the metadata are related to *TRL 1: Basic principles observed*, with 544 records (34%), followed by the *TRL 2: Technology concept formulated* with 29 %. The *TRL 3: Experimental proof of concept* (240) also have a relative high number of records (15%). The remaining levels (*TRL 4 - TRL 8*) made up 22% of the 1590 EIGR records.

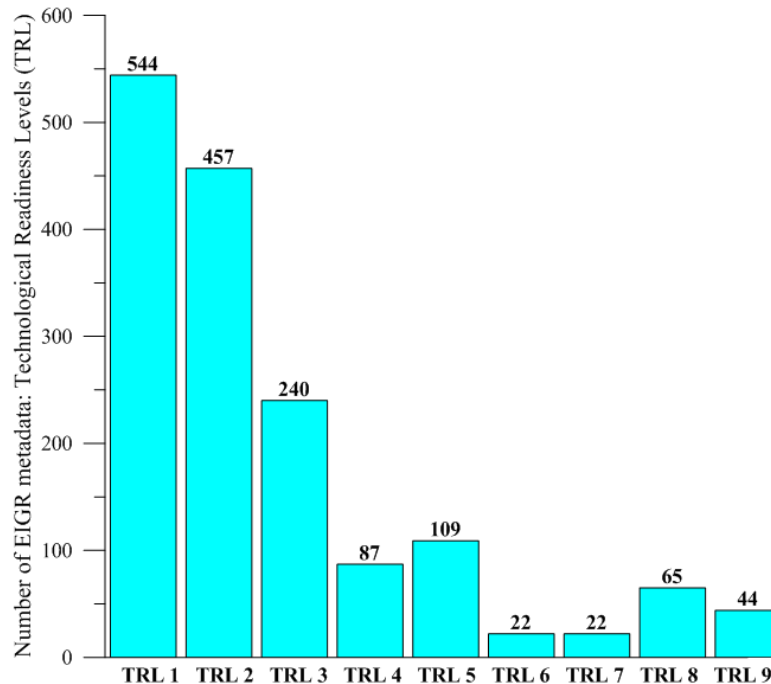


Figure 22. The distribution of EIGR metadata related to Technological Readiness level (TRL)

5 SUMMARY AND CONCLUSIONS

- The EFG Linked Third Parties (LTPs) representing 20 European countries have uploaded entries to the EIGR and 19 of these have provided Country reports.
- The Country reports involve answers for 6 questions, including information about the overview of the uploaded metadata, time consumed for uploading to EIGR, the main source of information used during the data collection, classification of their information based on the “Research and knowledge classes”, classification of their entries into 7 data/resource types (database, maps, books, etc.), barriers to find data and information.
- The EIGR metadata were inserted by the LTPs and the KINDRA consortium partners. In the statistics, data uploaded until 28/02/2018 were considered. Of the 2178 records 1968 EIGR metadata were uploaded by the LTPs, but the total number of metadata has been increasing. The LTP’s inserted their metadata and provided country reports until the end of February 2017, but they are improved their metadata especially for the translation of titles and abstracts. This additional work of the LTPs increased the amount of entrances complying with the standards by more than 10% until the end of December 2017. All these data will be taken into account in the gap analysis (WP3).
- Based on the LTPs information in the country reports, it appears that the barriers to find the data and the sources of information are different for country by country. The most frequent barriers were the language (most of the relevant information is in national languages) and the confidential-copyright issues, the main barriers were the scattering of huge amount of information. However, many LTPs didn’t report barriers.
- Nearly half of the EIGR entries were “National and local reports containing facts and data”, these entries were classified to Class 4 in the Research and Knowledge classes.
- More than the half (53 %) of EIGR metadata were assigned to the Operational Actions (OA) “Assessment and management” category of the HRC-SYS classification system. From the Research topics (RT) “Geology” topic was the dominant category, with ~ 76% of the total number of records. In the Societal challenges (SC) category, the majority of the metadata (~ 87%) belongs to the “Climate, environment and resources” category.
- Three EIGR visualization tools were developed and applied for detailed graphical evaluation and co-occurrence analysis of the large amount of information stored in EIGR database. To support and extend these analyses an additional external visualization tools (VOSviewer) were also applied.

- Based on the EIGR 2D category charts, a preliminary co-occurrence analysis was performed. It showed that the majority of the EIGR records were assigned to societal challenge no. 4 „Climate, environment and resources”; Operational Action no. 5 „Assessment and management” and Research Topic no. 4 „Geology”. Societal challenge no. 4 „Climate, environment and resources” and Research Topic no. 4 „Geology” were also receiving the largest number of entries for all of the other Operational actions (Mapping, Monitoring, Modeling, Water supply), For many intersections (category combinations) (43 intersections from the possible 125) of the HRC-SYS categories, “no entries” indicate areas with potential research gaps.
- The developed EIGR tools, and the external tools allow the further detailed analysis and visualization of EIGR metadata. The preliminary assessment shows that the features are suitable for the gap analysis pursued in WP3, which is one of the main purposes of KINDRA project. Consequently, a complete analysis of the EIGR content has been developed in D3.3.

ANNEX 1.

THE TEMPLATE OF COUNTRY REPORTS



KINDRA NATIONAL REPORT

Country/Association

Author:

** KINDRA has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 642047.*

1. Introduction

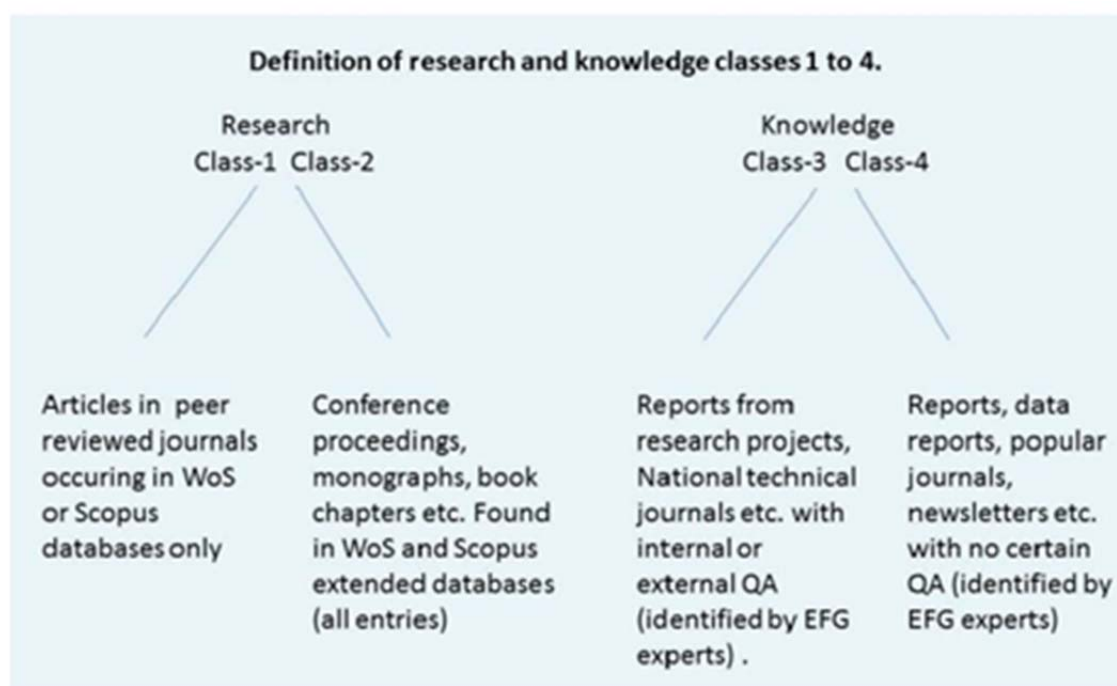
Summary on the total of data fill in the EIRG and time consumed.

2. Source of information

- 2.1. Institutions dealing with groundwater research/survey
- 2.2. Groundwater monitoring, availability of data
- 2.3. Journals/archives focused on hydrogeology

Please indicate here the main source of information used during the data selection

3. Type of information



Please indicate here how you judge the info to belong to one of the classes (class 1-4)

4. Topics

Please can you indicate how many input in the inventory you have for each of this categories:

- 4.1. National databases
- 4.2. National and local reports containing facts and data
- 4.3. Hydrogeological maps
- 4.4. Technical reports, Guidelines, Manuals, etc.
- 4.5. Books and book chapters
- 4.6. Position Papers and/or important papers on PR Journals
- 4.7. Others

5. Barriers to find data

Please can you indicate here barriers for find such data, for example: confidential, copyright issues, language, etc

6. Gaps in finding the information

To finish the report we would like to know you opinion on gaps in finding the information and perhaps suggestions on what to do about it.