

December 2020

DATALIT project

Needs analysis report - Executive summary (English)

In this executive summary we present the Needs Analysis Report introduction, where the aims and objectives of the DATALIT project are explained and what each chapter of the report contains. The main conclusions from the WP1 tools desk research, questionnaires, interviews and focus groups are also in this executive summary. Also, it represents the WP1 needs analysis report overall conclusion. This executive summary will be provided in 7 languages, in English and in the consortium partners native languages.

Introduction

This report is produced in the framework of the KA2: Cooperation for innovation and the exchange of good practices - Knowledge Alliance. This joint research project DATALIT, is under the Erasmus+ Program by European Commission, coordinated by the "*Istituto per le Tecnologie Didattiche*" of the Italian National Research Council and aim at promoting "Data Literacy" (DL) at the interface of higher education and business. The consortium has 14 partners, from 6 European countries, and it is envisaged to last for **2 years**, with a total budget of € 595.290,00.

The project DATALIT aims at:

- Offering university students across different Faculties, the opportunity to reach a suitable level of data knowledge that they can use and apply appropriately and diversely throughout their personal and professional lives;
- Narrowing the gap between business and academia and putting EU enterprises in direct contact with potential future employees or business partners properly trained to exploit data effectively.

The overall objectives of the project derive from the challenges and needs identified during the preliminary need analysis, and are the following:

- Defining a common understanding of what a data literate person is supposed to know and is able to do, not only as worker but also as an active citizen.

Project reference no. 612561-EPP-1-2019-1-IT-EPPKA2-KA

- Designing and developing innovative didactic frameworks related to data literacy that reflect the views and needs of academia and job market.
- Fostering a culture of data literacy among European citizens
- Making graduate students acquire the data skills they need to success in the job market and to actively take part in the civil society.

This report reflects the deliverable for Work Package 1 (WP1) - Developing a comparative needs analysis on Data Literacy, for the deliverable Needs analysis Report - that compiles the work of the 14 consortium partners. The needs analysis report summarizes the results of the research activities carried out in the WP1. This report will be divided in two main parts: the first one summarizes and enhance the insights from the desk research. The second part will present the results obtained by conducting the interviews, questionnaires and focus groups.

In Chapter 2 the first part of the WP1 deliverable Need Analysis Tools compiles the Desk Research conducted by all DATALIT partners for **19 European countries** and Tunisia. Chapter 3 consists of the second part of the WP1 deliverable Need Analysis Tools, with three tools used for the analysis, the Questionnaires, Interviews and Focus Group. Finally, Chapter 4 is the conclusion of this report, with the major findings and conclusions of all the deliverables presented, thus finalizing the last WP1 deliverable Needs analysis Report.

1. Desk Research

The desk research consists in the analysis of different sources, by all partners of the DATALIT consortium, for **20 countries**. The multidisciplinary approach of this research allows us to observe how data literacy status is in each country. Although, some countries provide more information than others, we could still retrieve important information, which main conclusions are as follows.

The term ‘data literacy’ isn’t well known in most of the countries analysed. The most widely used terms are ‘digital literacy’, ‘information literacy’, ‘data competence’, ‘media literacy’, ‘statistical literacy’, ‘computer/IT literacy’, among others. In most countries is closely related to digital skills. Some countries, like Belgium or Serbia, don’t even recognize or have a direct translation of the term itself. Although, all countries agree that ‘data’ is very important and how people use, analyse, and perceive that data is essential for today’s job market and decision making.

Project reference no. 612561-EPP-1-2019-1-IT-EPPKA2-KA

Every desk research agrees with the existence of different stakeholders related to DL, namely individuals, business, universities, government, media, NGOs, research institutes, and others (i.e. all those who work with data). Relative to DL stakeholders' statistics, in the different countries, the information is scarce. For example, in Lithuania there is a vast and various information about DL stakeholders; in Portugal there is statistical information about digital literacy stakeholders, not DL; and other countries have little to no statistics on DL stakeholders.

Universities are still the main providers of DL related courses, but there are other organizations that also provide DL courses or DL related courses. Each country has specific private companies/NGOs/government entities, that provide these courses, like, for example, DigitYser is the Digital Innovation Hub of Brussels (Belgium), or Digital Serbia Initiative (Serbia). At an international level, there are a few companies that deliver specific DL courses/programs/certifications and are the greater players in this field, namely Qlik (with the 'Data Literacy Program'), Microsoft (with the 'Microsoft Data Science' program), Data Literacy Project (from Qlik, with several courses on DL), and online providers, like Coursera, edX, Udemy, etc. Learning Management Systems are widely used in most countries, especially Moodle, also e-portfolio Mahara, and other proprietary tools.

In some countries there is support from public and private institutions for the acquisition and validation of DL related competences, like IEFP in Portugal, or the Department of work in Belgium, for example.

In HEIs/Universities, there is a great range of DL related degrees (bachelor, masters, PhD, and others), that include courses in data science, big data, business intelligence, artificial intelligence, and other IT related subjects, but there seems to be an increase in interest on DL related subjects in non-IT degrees, like marketing, tourism, journalism, social sciences, etc. In HEIs we can see that DL is always part of a degree, or course, not exactly the whole course. Many universities have partnerships with private companies/institutions, that in one way or the other influence the path of these HEIs courses, and help with students' internships, mentoring and projects.

In the enterprise domain, several job websites were analysed, and the research shows that for the specific term 'Data Literacy' there isn't many offers, but when it comes for the terms data, data science, big data, AI, and other IT related subjects, there is a great demand

Project reference no. 612561-EPP-1-2019-1-IT-EPPKA2-KA

for professionals that are data literate, and have great programming skills. Other offers, like administrative roles, HR, accountant/finance roles, also demand professionals with some sort of data literacy, like managing databases, using MS Office, and other kind of programs/software/tools. Many different hard skills are asked in these job announcements, but overall the soft skills are more or less the same: motivation, flexibility, leadership, team player, problem solving capacity, proactive attitude, strong interpersonal skills and cross-cultural competence, empathy, attention to details, communication skills, independent worker, English fluency, etc. It seems to be a rise in demand for professionals that know how to work with data.

In terms of DL related competences and validation many of the countries in this research don't have a validation approach to DL, other than formal education. Most validation processes include certifications delivered by private institutions/NGOs (e.g. The Data Literacy Project), ECTS delivered in universities, and certifications delivered by government/public entities (e.g. Portuguese Agency for Qualifications (ANQEP), in Portugal for competences related to DL, but not DL exactly). Most DL competences are recognized in the business sector and internally validated, but there isn't quite a structured system/tool or framework to assess these competences for the general public. Some countries (e.g. Serbia) don't have institutes or organizations, involved in working on Validation of Informal and Non-Formal Learning (VINFL), others (e.g. Spain) use European frameworks for this, like DigiComp 2.0.

The main conclusions in this desk research are that there is a lack of understanding of what 'Data Literacy' is and what a data literate person knows and does. There is a clear confusion between different kinds of literacy, specially between digital literacy and data literacy. In many countries data literacy = digital literacy, and this is a false statement. It becomes necessary to clarify the meaning of each type of literacy - data, digital, information, media, statistical, and others - and how they interact with each other, because none of them is totally independent of the other. In every country there is a qualification system or framework for different DL related competences, but specifically for DL this is non-existent. DL involves different soft skills, that are appreciated and recognized by the business sector and are validated in an internal level, but there isn't a way to assess these skills for the general public or validate them. Although there are private entities/NGOs that offer different types of validation and assessment and some governments are involved in different projects with this in mind, there a shortage of tools

Project reference no. 612561-EPP-1-2019-1-IT-EPPKA2-KA

and frameworks to properly do this at a national level. ‘Data literacy’ is a term that is not well known, but everyone uses DL competences/skills, within those who work with data, so a strong structure of the concepts and of the technical and non-technical skills is needed to help identify a data literate person and create more competent professionals.

2. Questionnaires

The objective of the online questionnaire was to fulfil the WP1 - Developing a comparative needs analysis on Data Literacy - project deliverable 1.2, regarding the Needs Analysis Tools. It was expected to have a total of 500 participants in this online questionnaire. The online questionnaire had a total number of respondents of 705, but the number of usable answers is 704. All consortium partners participated and the numbers of respondents per partner country are:

- EN = 48
- AL = 73
- LT = 75
- SRB = 179
- PT = 198
- IT = 128
- DE = 3

The questionnaire was conducted by Qualtrics¹ platform, using surveys in seven different languages, English, Albanian, Lithuanian, Serbian, Portuguese, Italian, and German. The questionnaire was disseminated, mainly, via each partner organization (internally) and social media.

In the first section the respondent’s demographics were analysed. The majority of respondents ($\approx 74\%$) have 40 years old or below and, approximately, 60% of them are male. Respondents from 27 countries participated in the questionnaire, which the main contributors were Portugal, Serbia, Italy, Lithuania, and Albania. 55.4% work in companies and 20.2% in HEI, while the remaining respondents work at training providers, public entities, NGOs, and other kinds of organizations.

¹ <https://www.qualtrics.com>

Project reference no. 612561-EPP-1-2019-1-IT-EPPKA2-KA

The second section refers to DL in general. 54.2% of respondents rated themselves with medium knowledge in DL. Approximately, 7% of respondents admitted they don't have any knowledge in DL. 64.1% agree that personal development is the most important feature for DL. The four most important soft skills to have are evaluating/reflecting, critical thinking, problem-solving, and communication. 7.6% agree that leadership is a soft skill unimportant to DL. In terms of functional competences, reading/creating data classification rules is the most important, according to 56.2% of respondents.

The third section analyses the acquisition and development of DL related competences at the respondent's organizations. 48.4% have a moderate promotion, by the organization, of acquisition and development of this competences, 25% have a rather strong promotion to acquire and develop this competences, and, approximately, 8% admit they don't have any help from the organization in this matter. Unfortunately, between 33 and 59% of respondents don't use any kind of tool or method to acquire and develop DL competences. Approximately, 64% of respondents say that their organization does not use any kind of validation, for DL related competences. The remaining respondents say the organization uses tools like LEVEL5, DigiComp, ESCO, job evaluation, tests and surveys, certifications, etc. The respondents have used frameworks for validation, like ESCO, DigiComp, EQF, and others. In their organizations, they often use online tools and use observation of facilitators/mentors, as a way to assess DL related competences. In terms of European frameworks for validating DL related competences, between 77% and 90% say they don't know or don't use any. Among the respondents that answered yes, the most known/used European framework is the 'Key Competence Framework for Lifelong Learning'. About 69% of respondents use e-learning and/or blended learning in their organization, where they mostly use videoconference and LMS digital tools, while the least used tools are e-portfolio and MOOCs. Approximately, 72% of respondents don't know any open learning systems that connect with validation tools.

On the overall, it is possible to say that there is still some lack of knowledge in data literacy. Analysing the respondent's answers and suggestions, it is clear that there is some confusion regarding DL soft skills and DL functional competences. In the text boxes, where respondents could put other soft skills they considered important, some of them wrote down technical skills, like 'Digital skills' and 'Statistical knowledge', instead. The same occurred in the functional competences text boxes, where respondents wrote software/tools instead. It becomes necessary to clarify these definitions. When it comes

Project reference no. 612561-EPP-1-2019-1-IT-EPPKA2-KA

to the acquisition and development of DL related competences, it is clear that organizations don't properly promote this acquisition and development, by not using tools and methods for this. And, in terms of validation of these competences, most organizations don't use validation tools, while most respondents know or used this validation tools/frameworks. To assess DL related competences, organizations don't use all the digital tools available for this, especially e-portfolio and MOOCs (that are normally free). Most respondents, said that that they don't know open learning platforms connected to validation tools, the remaining respondents mention many online learning platforms for this.

In conclusion, we can say that this questionnaire brought to light issues regarding the lack of knowledge on validation frameworks/tools in DL related competences. Many respondents work in IT roles and similar, and don't know/use validation frameworks to validate digital skills (e.g. DigiComp). Other issue regards the digital tools used or little used by organizations, there is a need to create some awareness on the existing tools to acquire and develop DL related competences, so that organizations may use them properly and promote them among their employees/students and general population.

3. Interviews/Focus groups

This report represent the work of the DATALIT Consortium partners by completing the last two deliverables from WP1 - Interviews and Focus Group. The partners did between 1 to 4 interviews each, to different professional experts, to understand how these professionals perceive the need of DL in their countries. The partners also conducted one focus groups each, except for UET that did two focus groups, one for the business sector and another for the HE sector. The focus groups had different professionals from HE or Business sector, or both, and they discussed their opinions, perceptions and organizational issues related to DL, in their respective organization and country. Due to Covid-19 restrictions, most of these interviews and focus groups activities were conducted online.

The interviews were conducted by each partner and had the inputs of various experts. Overall, the different interviews align with the same conclusion, "Data Literacy" isn't a term well known by all. Many interviewees relate DL with digital literacy, data science, and other technical areas. The demand from companies and academia for a data literate professional is high (depends on the field of work), but there aren't many learning offers in this area of expertise. The HE interviewees admit that DL is present in different study

Project reference no. 612561-EPP-1-2019-1-IT-EPPKA2-KA

programs at their organization, but many say that they don't know any specific offer for DL alone. In companies the learning offers are limited, mainly by online courses or informal on-the-job training. Although, some companies have created their own online learning platforms, so that employees can choose any learning path at no cost (e.g. Salvador Caetano Academy, Inova+ interview). Stakeholders are normally engaged in DL activities, mainly at HEI, but not all of them (e.g. University of Palermo, CNR-ITD interview). Young graduates have the capacity to work in most of the interviewees' organizations, but only if they come from technical HEI's, where data skills and knowledge are the main study field. Other HEI's have a lack of data related learning offers and the young graduates may not have enough literacy on data. All the interviewees refer that there are important skills necessary to learn and work with data, like critical thinking, communication, learning to learn, problem-solving, evaluating/reflecting, flexibility/adaptability, networking, research, management, client orientation, IT skills and others.

The focus groups were conducted by each consortium partner in their organizations or online (due to Covid-19 restrictions). The focus groups consisted in a discussion group with 4 to 10 professionals each. These professionals were from higher education and/or business sectors, each from one's area of expertise, and helped DATALIT bring to light the issues related to DL in their respective organizations and countries. Overall, the participants of the different groups agreed that DL can be taught, it has to be trained, by using practical cases and other methods, because DL is necessary for one's future professional career. Several subjects were discussed, but some findings were unexpected like, for example, many discussion groups refer that DL competences should be taught at lower levels of education, from primary to secondary schools, so that this competences are present in students before they reach university. Other findings revealed that DL isn't a well-known concept and many participants agreed that there aren't any specific DL training/learning opportunities. Most learning of DL related subjects are specific to data science, digital literacy, big data, statistics, and other data related fields, not DL itself. Some groups discussed the need for certification/validation of DL competences. Although, many of the participants work daily with data, the existing European validation frameworks aren't known/used by many.

Through the two tools (interviews and focus groups) used for this report we can sort out some interesting findings:

Project reference no. 612561-EPP-1-2019-1-IT-EPPKA2-KA

- “Data Literacy” isn’t a known concept in general.
- It becomes necessary to create a broader DL concept that can reach a bigger audience.
- There aren’t any DL specific courses/training/learning activities, only DL related.
- There should be more practical learning opportunities (e.g. case studies).
- The knowledge/competences on DL should be taught at lower levels of education.
- The learning of DL should be adapted to the target audience - teaching DL to a retired citizen is different from teaching it to a university student, for example.
- There is a need to create a validation framework, specific to DL, that can be known and used by all.

Overall, DL is still a new and confusing concept for organizations and individuals, therefore the work of the DATALIT project becomes even more important.

Conclusion

In this chapter we present the overall conclusions for the needs analysis report. In each chapter we presented the summaries and main conclusions from the WP1 tools used. So, in this conclusion we will focus only on the obtained results and main points.

Desk Research

From the 20 countries analysed, one of the main conclusions that the desk research provides is that there is a lack of understanding of what ‘Data Literacy’ is and what a data literate person knows and does. There are different types of literacy (digital, information, etc.), and many are confused as DL, because there isn’t a clear definition of DL in many countries. For DL, there isn’t a framework or qualification system in any of these countries, only for DL related competences (e.g. digital literacy). Soft skills/competences in DL are appreciated in the business sector, but these are normally validated informally and internally. So, it becomes necessary to define what DL is, create a framework capable of validate a data literate person, and in this way help create professionals with stronger DL competences and knowledge.

Questionnaires

The online questionnaire was an important tool to analyse what the general public think and knows about this subject. The questionnaire reinforces the desk research findings about the lack of understanding of what DL is. Many respondents have, also, some

Project reference no. 612561-EPP-1-2019-1-IT-EPPKA2-KA

confusion about the difference of soft skills/functional competences and technical skills.

The questionnaire brought to light issues regarding the lack of knowledge on validation frameworks/tools in DL related competences. Many respondents work in IT roles and similar, and don't know/use validation frameworks to validate digital skills (e.g. DigiComp). Organizations don't properly use, or use very little, the available digital tools, according to the respondents. So, there is a need to create some awareness on the existing tools to acquire and develop DL related competences, so that organizations may use them properly and promote them among their employees/students and general population.

Interviews/Focus Group

The interviews and focus group were the last two tools used in WP1. The experts interviews were conducted individually, by the consortium partners, to assess how these professionals perceive the need of DL in their countries. The focus groups were done with groups of experts from HE and business sectors, where they discussed their opinions, perceptions and organizational issues related to DL, in their respective organization and country. Due to Covid-19 restrictions, most of these interviews and focus groups activities were conducted online.

As mentioned in the previous chapter, these two tools revealed some interesting findings:

- “Data Literacy” isn't a known concept in general.
- It becomes necessary to create a broader DL concept that can reach a bigger audience.
- There aren't any DL specific courses/training/learning activities, only DL related.
- There should be more practical learning opportunities (e.g. case studies).
- The knowledge/competences on DL should be taught at lower levels of education.
- The learning of DL should be adapted to the target audience - teaching DL to a retired citizen is different from teaching it to a university student, for example.
- There is a need to create a validation framework, specific to DL, that can be known and used by all.

These findings, along with the results of the desk research and the online questionnaires, makes us conclude that DL is still a new and confusing concept for organizations and individuals, that there isn't a validation framework for DL competences and that “Data

Project reference no. 612561-EPP-1-2019-1-IT-EPPKA2-KA

Literacy” is essential for all of the professional that work with data. Therefore the work of the DATALIT project becomes even more important.

