DATA LITERACY: FROM HIGHER EDUCATION TO THE WORKPLACE

Johanna Walker, Manuel Leon-Urrutia, Sarah Hewitt, Davide Taibi, Vera Pospelova, Sergio Splendore, Andrea Nelson Mauro, Tim Scholze, Laimute Urbsiene, & Ugljesa Marjanovic

https://dedalus.pa.itd.cnr.it/en/

DEDALUS is a joint research project funded under the Erasmus+ Programme. DEDALUS is funded under Call KA203-INDIRE (Italian Agency for E+). Project number: 2019-1-IT02-KA203-063359.

DEDALUS

Table of Contents

DEDALUS

Executive Summary

> According to the DEDALUS project, "Data literacy is the ability to read, write, critically assess, and communicate data in context, including an understanding of data sources and constructs, analytical methods and techniques applied — and the ability to describe the use case, application, resulting value, and its implications."

In this report, through desk research we review how data literacy modules and programmes are currently delivered in higher education (HE). We then interview 16 senior employees in small and medium enterprises (SMEs) across Europe to discover the importance of data literacy in their organisations, and their views on how well employees are equipped for the data demands of the contemporary workplace.

We distinguish between four types of broad data literacies. Basic digital skills are often learnt in school/first job, and basic data literacy is often accumulated via job experience. Advanced data skills are often learnt through HE and job experience but advanced data literacy skills are more complex to acquire, however, this could be addressed in Higher Education or via on the job training (continuing professional development).

Ultimately, we identify three key actionable insights for data literacy programmes in HE to support SMEs in being able to access a suitably data literate workforce

1. Ensure the ***provision of the full spectrum of data literacy***, and not just data skills;

2. Build on generic computer science modules with ***bespoke, relevant elements for specific fields and professions***, to help graduates develop the ability to extract value from data insights and application earlier in their careers;

3. Develop ways to ***evidence and certify the knowledge and insight*** aspects of data literacy, in order that employers are more able to identify and recruit for these skills.

Glossary and acronyms

HE	Higher Education
EQF	European Qualifications Framework
Higher Education	EQF Levels 6, 7 & 8 Education
CPD	Continuing Professional Development
SME	Small and Medium Enterprise
DigComp	European Digital Competence Framework for Citizens
SaaS	Software as a Service
CRM	Customer Relationship Management

Introduction

DEDALUS aims to DEvelop DAta Literacy courses for University Students. This will equip students with the necessary competences to cope with future digital challenges and to create an additional value for the enterprises and industries where they will be employed. At the heart of DEDALUS is an innovative modular, open and online learning curriculum to include data literacy competences in any study field, and a CPD qualification for Higher Education professionals, enabling them to transfer these competences in their practice.

The need for these data literacy courses is to ensure European enterprises can recruit potential future employees or business partners properly trained to understand and exploit data effectively. Doing this requires a sound stocktaking on successful approaches on how to include data literacy in university programmes, and a thorough understanding of the unmet data training needs of the small to medium business sector. Together, this creates a framework of data literacy competencies that are appropriate for business success in Europe's Digital Decade[1]. This framework will inform the devising of an innovative, modular, open and online learning curriculum to equip students with the necessary competences to address future digital challenges and to create additional value for the enterprises and industries where they will be employed.

This report shares our findings about existing provision, our insights into small and medium enterprise (SME) needs, and combines these to develop a framework of competences.

The unmet need

Small and medium enterprises we spoke to told us they recognised data literacy as an *up and coming area* and a *minimum of basic skills is necessary*. Yet they also said that *universities are not teaching data literacy to the required standard* forcing people to learn on the job, and that *costs and barriers to SMEs of training are high.*

On the other hand some companies were concerned that *a focus on data literacy should not be too onerous* for companies. An example they gave was that of the GDPR, which *many small companies find challenging* and difficult to comply with [2]. This has made some digitisation aspects arduous for small companies.

There is also an awareness that data literacy is still to some extent undefined with low awareness in the general population. The concept of data itself is still opaque to many, who think of it in terms of *0s and 1s*, and being digitally competent or data literate is still frequently equated with being capable with packages such as MS Office. Data literacy itself still has *unknown unknowns.* Yet, in our post-truth, deep-fake world, there is increasingly, *a civic responsibility of organisations to communicate data-based information clearly"*, which urgently requires that employees should be appropriately educated.

[1] https://ec.europa.eu/commission/presscorner/detail/en/ip_21_983

Who should read this?

This report is aimed primarily at two audiences. The first is SMEs who are keen to develop data literacy within their organisations, and want to know what that might look like. The second is educators in higher education institutions, who are developing curricula to support data literacy across multiple industries and roles.

In this report

It is key to establish exactly what we mean by data literacy. We briefly explore dimensions of data literacy found in research, and then create our own definition. We expand on this to describe what practical skills this might imply in the workplace. We then survey how data literacy is taught in higher education institutes across Europe, and map these offerings on to our definition.

In the second part of the report we interview representatives of 15 SMEs across Europe to discover how they understand data literacy, the importance of it to their organisations and whether they feel their workforce has the appropriate data literacy skills.

Based on this we create a competencies framework. We then perform a gap analysis between this desired framework and the courses offered in higher education institutions, to identify the curriculum gap.

"I think that a lot of people who don't have
a science or computing background
don't really recognise what data is.
They can't point at something and go 'that is data'."
Arts and Culture Organisation CEO

How this report was created

Our research looks at two areas - what universities currently make available in terms of data literacy education, and what the needs of SMEs in Europe might be in terms of a data literate workforce and training. Based on this we create a gap analysis to identify unmet needs.

To identify current data literacy provision, systematic desk research on higher education institutes in 14 European countries was performed. The observed countries were Portugal, France, Austria, Latvia, Netherlands, Croatia, Romania, Ireland, Spain, Italy, Serbia, Germany, the UK and Romania.[1]

We addressed the following questions:

Q: How is data literacy approached in the Higher Education competency frameworks in your country-region?
Q: How do Higher Education curricula across all disciplines cover data literacy related topics?
Q: Do national HE curricula address specific aspects of data literacy reflected in national/European digital competence frameworks (e.g. DigComp)?

To canvas views on the current level of data literacy of employees across Europe, we conducted interviews with senior representatives of several small and medium enterprises across Europe. These organisations were in the medical, business services, consulting, humanitarian, e-learning, digital marketing, fintech, IT and arts and culture fields. The observed countries were Italy, the UK, Spain, Germany, Austria, Ireland, Portugal, Serbia, Croatia, Lithuania, Latvia, and Romania.

We asked them the following questions:

Q: How would you describe a data literate person?
Q: How important is the ability to understand, analyse and express data for your company?
Q: How many of your employees you would say are data literate?
Q: Are you satisfied with the data literacy competences of your staff?
Q: Do you think that your company would benefit from employees with stronger data competences? How?
Q: What kind of data related competences do you look for when hiring a new employee?
Q: What other soft competences (personal, social and organisational) a new employee should have?
Q: Where do you think they should ideally have acquired these competencies (university, previous experience, cpd training)? [3]

Finally, the background to this report is informed by multiple other publications and articles, which can be found in the ‘Further Reading’ section.

What is data literacy?

"We are all illiterate when the text is in a language we don't know and we have no dictionary."[2]

Wildly incomprehensible statistics about the sheer volume of data in our digital world abound. Regardless of whether there is 44 zettabytes (10 to the power of 21 bytes) of data on the Internet, or whether so much data is generated each second that it is the equivalent of 1.7mb for each person in the world, we live with a potential abundance of data which, appropriately used, can provide opportunities for economic growth, social change, sustainable development, transparency and engagement. By itself, however, data is insufficient. To be useful, people need to be able to extract information from it, and understand what that information means. The ability to do this - to be data literate - is rapidly becoming a fundamental requirement to participate in contemporary life. [4]

Exactly what constitutes data literacy has been much researched and discussed in recent years. Despite this, data literacy definitions vary, depending on the discipline or country. Linguistic diversity in Europe means that in some countries, such as Spain, Serbia and Lithuania, there is actually no direct translation.

The European Digital Competence Framework for Citizens, also known as DigComp and first created in 2013, offers a tool to improve citizens' digital competence. DigComp has 5 sections, the first of which is "Information and data literacy." They describe this as being able, *"to articulate information needs, to locate and retrieve digital data, information and content. To judge the relevance of the source and its content. To store, manage, and organise digital data, information and content.*"

Researchers have suggested that in fact data literacy is more complicated than this catalogue of competencies captures. Data literacy may vary in context, and it is, *"important to consider the needs of a particular community before prescribing a data literacy curriculum on a broader scale"* [5]. In addition to similar aspects of data literacy as outlined in DigComp, other researchers extend the definition to include, *"consideration of*

[2] https://openjournals.uwaterloo.ca/index.php/JoCI/article/view/3274/4297

ethical use of data". They also note that, *"[data literacy] is based on core practical and creative skills, with the ability to extend knowledge of specialist data handling skills according to goals."* [6]

Argast and Zvyagintseva [5] see variation horizontally - that the varying needs of different communities will affect the specifics of the literacy; while Woolf et al [6] see the variation more vertically, noting that more specialist roles will require greater data expertise. Ultimately, the required knowledge, skills and attitudes related to it are dependent on the context; as is the competence level. Woolf et al also introduce the idea of data literacy as compassing an ethical awareness. The Open Data Institute suggests that, much like conventional literacy, there is a skills side, comparative to language, and there is an interpretative side, comparative to literature.

In the DEDALUS project we acknowledge these additions. Our definition is,

"Data literacy is the ability to read, write, critically assess, and communicate data in context, including an understanding of data sources and constructs, analytical methods and techniques applied — and the ability to describe the use case, application, resulting value, and its implications."

The term implications is vital, addressing, for example, how to deal with privacy, how to deal with security, the potential consequences if data is not handled correctly and other ethical, legal and regulatory aspects.

While the 6 key aspects of our definition are applicable to all forms of data literacy in a variety of domains, our more in depth explanations are focused on the general business sphere.

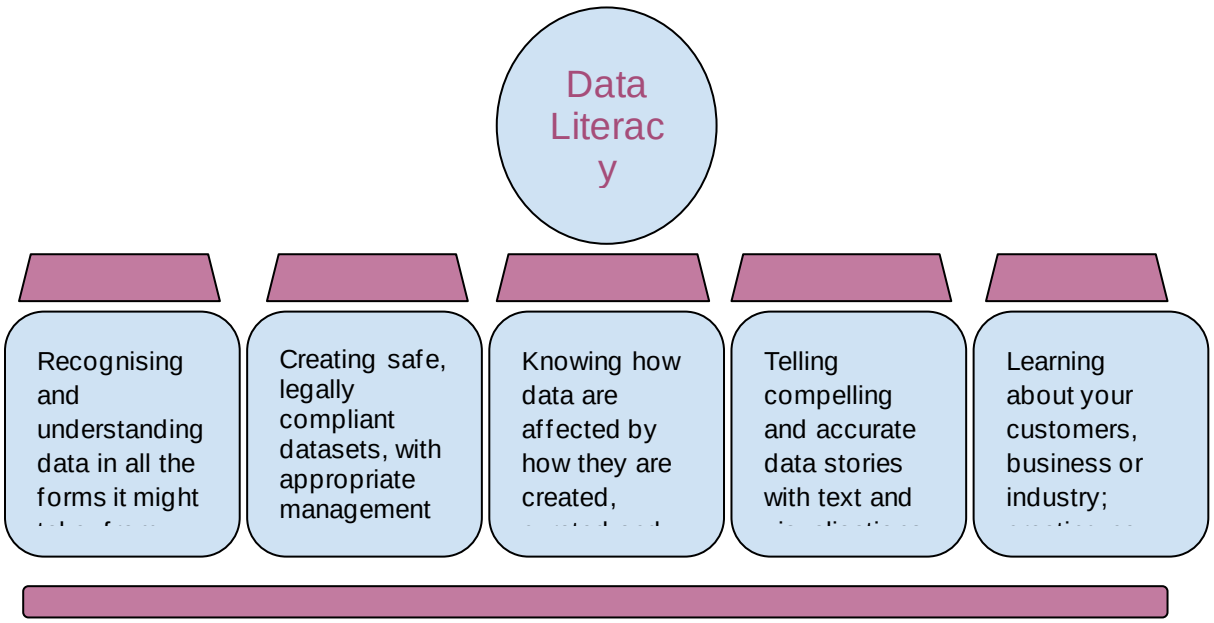

Figure 1: The Dedalus definition of data literacy and its practical application in the business sphere

"From the current market trends, [data literacy] will be an increasingly important skill in the future especially if it is also integrated with specific skills in the [relevant industry] sector."
Digital Marketing Director

How is data literacy currently taught in higher education?

Currently these are largely taught in higher education institutions across Europe by faculties across academic disciplines including computer science modules that address data literacy aspects in their offering.

For example, business faculties across most observed universities included modules on Big Data, data structures, cloud computing, data analytics and visualisation.

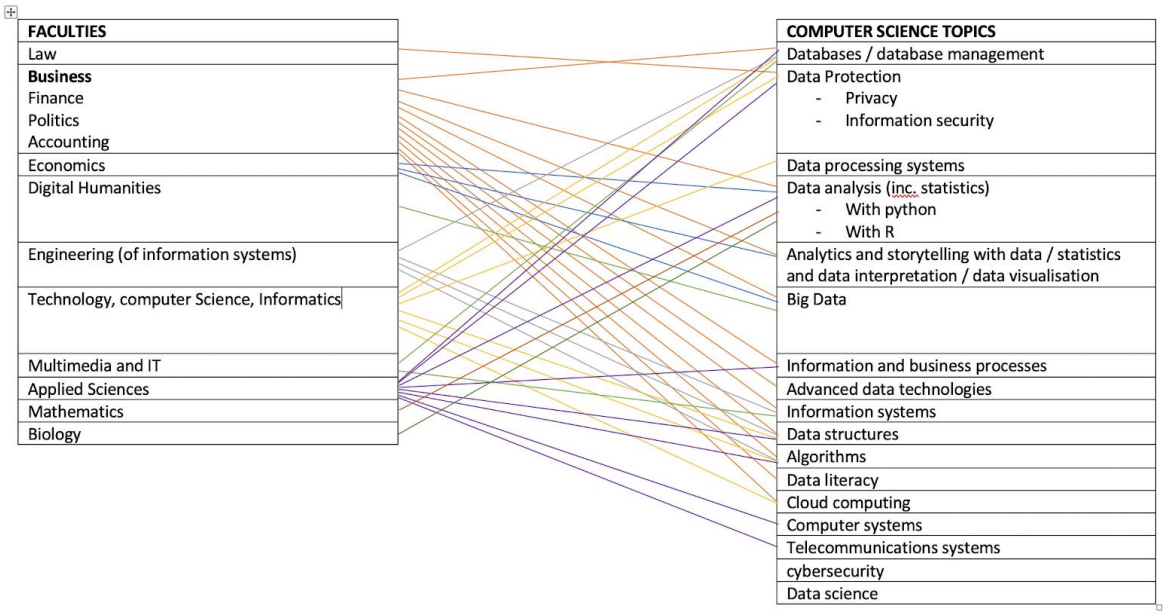

Figure 2: Students in the university faculties on the left are frequently offered modules in the computer science topics linked on the right.

These programmes are aimed at developing a wide range of competencies. After curating all these competencies, five key clusters of competencies can be identified:

- Data protection and security: Protecting devices, content, personal data and privacy in digital environments complying with existing laws, ethical standards and best practices.
- Data selection and critical assessment: To recognize and select relevant data for identified needs, in order to extract value from it.
- Data processing: To identify and select the most effective data processing and pre-processing methods using state-of-the-art tools and data formats, to ensure and enhance the quality of data.
- Data analysis: To use both quantitative and qualitative methods to extract insights from data in different formats from different sources.

- Data visualisation and storytelling: Apply best practices of data visualisation in specific contexts (business, science, academia) by using relevant tools and methods, to effectively convey analysis results and create compelling narratives.

Competency Cluster	Related DEDALUS Data Literacy Concept
Data Protection and Security	Write
	Extract Value
	Understand Implications
Data Selection and Critical Assessment	Critically Assess
	Extract Value
	Understand Implications
Data Processing	Extract Value
Data Analysis	Extract Value
	Communicate
Data Visualisation	Read
	Communicate
	Critically Assess

Table 1: Mapping the identified competency clusters provided by universities across Europe to the DEDALUS data literacy definition

DEDALUS

"Analysis can be a commodity,
the strategic capacity [to exploit insight] cannot"
Communications Consultancy Manager

What data literacy skills do small and medium enterprises need?

How do employers define data literacy?

Unsurprisingly, most of the organisations we spoke to mentioned analysis as a key data literacy skill, but a number also noted that, *data analysis is a [only] means to achieve an objective,* in other words, to develop actionable insights, to enable staff to make decisions, or to enhance weak signals from the market or customers.

What was particularly interesting was that when describing their understanding of data literacy as a concept many employers discussed issues of collection, curation, cleaning and critical and assessment of data. How to find the right type of data, how to compare sources, how to generate information from *unusable data"* how to prepare data and how to form an opinion about data were all cited. While this did not emerge frequently while discussing what kind of data literate tasks employees actually performed, it is clear that these types of activities - not particularly associated with any given role or software - are in employers' minds when they consider the concept of data literacy.

When asked to describe the competencies of data literacy, employers did not reference specific software packages. In Insight 3 (below) we see this has not yet fully translated to practical implementation in the workplace; when asked which data literacy competences they looked for when hiring, specific software packages were frequently enumerated.

Issues of privacy, compliant data handling, the ethical use of data and so on were not normally defined in the description, beyond one organisation which particularly focused on compliance.

Is data literacy important to SMEs?

Almost all our interviewees felt that data literacy was important, crucial, fundamental or key. This was true regardless of whether the context was plainly data heavy (such as a research or analytics environment) or less obviously so, such as an arts and culture business.

Data literacy was prized for both internal decisions, such as *cases in which fundraising strategies need to be decided based on revenue and cost projections* and external

interactions, *with the different actors involved in the company's relations.* Being able not only to perform or understand analysis, but also communicate the importance ot that analysis to funders, or clients, or the chief executive, is seen as vital.

Are SMEs satisfied with the data literacy of their staff?

Responses to this question varied. Some employers felt the overall standard was not strong enough, others felt sufficient staff were appropriately data literate to support the activities of the entire company and others felt very satisfied with the data literacy levels of all their staff.

However, most interviewees were of the opinion that data literacy is on a spectrum. Most organisations felt very happy with the data literacy of their most technical and/or senior staff, but less so with their more junior or generalist staff. One such comment was, *the data understanding and more importantly – data contextualisation and data-led project creating – is better among senior position employers within our company"* This fits with the idea that having the necessary background to exploit the data is important - this context will be accumulated over a longer career. Some interviewees believe that data literacy is not an immutable skill, noting that, *there is always something to update, because there are new techniques and new tools.*

There is a gap between the two activities of analysis and action that were identified earlier: *in terms of data mining, very literate, in terms of transforming it into analysis that Csuite can understand than probably less so.*

We need people who can work with data, drive the tools, write code and manipulate the data. But we also need people to be able to critique the use of data, understand its constraints and its impact on society. We need this blend of skills to get the best out of data.

Jeni Tennison, VP Open Data Institute[3]

INSIGHT 1: More technically skilled employees are perceived as more data literate...

Unsurprisingly, employers in more technical industries express greater satisfaction with their employees' level of data literacy, similarly, staff in more technical roles in non-technical industries are also perceived to be more generally data literate.

This suggests there is still an issue around the idea that someone who specifically works with data is 'data literate', rather than having fully assimilated the idea of a generic data literacy. However, a number of employers are aware that a need is opening up for more generic skills and this needs to be addressed.

Such generic skills are those around issues like communication and understanding; extracting meaning from visualisations and interpreting data. These are softer, less technical skills that are applicable across multiple roles, tools and activities. They are less easy to articulate, specify and assess on a job description.

INSIGHT 2: ...but general data literacy is required in the workplace

More data literate staff were perceived as helping the organisation in its environment, understanding the market and customers better, and being better able to deal with clients and suppliers. There was an understanding that data is not only about being online and digital, but is equally important for delivering offline products and services. A fairly typical comment was, *[Our] level of data literacy competence is satisfactory. However, if you ask me if we would like to see more data competent candidates, I would definitely say yes.*

There was an interesting dichotomy between specialisation and generalisation - sometimes present in the same employer - regarding whether it was preferable to have an extremely data literate person, a data expert in the company whom everyone could consult and be used as a resource, or whether enabling everyone to be autonomously capable in data was

[3] https://theodi.org/article/data-literacy-what-is-it-and-how-do-we-address-it-at-odi/

the right way to go. However, those who had very data literate employees who were used as a resource by less literate colleagues found this inefficient.

Having particularly skilled data specialists in an organisation of low data literacy levels also means that there is a gap between the commercial arm and the data people, which may lead to failure to exploit the data, *to create truly valuable activities, the ones that will initiate the economic growth.* As one respondent commented, *analysis can be a commodity, the strategic capacity cannot.*

INSIGHT 3: Employers still think mainly in terms of specific software skills when hiring

The level of data competence sought when recruiting is hugely dependent on job level, position and description. At an entry or junior level data literacy is functionally described as the ability to use software such as Office, Excel, or spreadsheets. At a more senior level employers described looking for skills such as big data; ability to merge different datasets; programming in R, Stat, Matlab; data mining; artificial intelligence; forecasting; customer relationship management; automation marketing; web and social media analysis; database management; Google Adwords; SPSS; data processing and visualisations. These can largely be characterised as pertaining to specific packages or job roles.

Some employers articulated that they also recruited for more general data literacy skills, such as the *ability to comprehend various data resources*, to *interpret the client needs from the data,* to have the *capacity of abstraction and generalization of the data, understanding of what questions can be answered by data* and *an understanding of possible data sources* and *not only knowing how to read them but to handle them responsibly.*

Even those with excellent technical skills may lack these more general data communication skills. As one employer told us, *the only thing I find slightly frustrating is trying to instill within the data engineers and or analysts how their work, their analysis might be viewed by someone with zero or very little technical understanding*.

There is some doubt that a data literacy course that could address the needs of everyone in an organisation could exist: *in my case this would not work, because of differences and specialization in our office. It will be tremendously hard to find the course that will benefit and will be interesting for majority of the office*.

INSIGHT 4: General data literacy can be usefully framed as a soft competence

The organisations we spoke to were experienced in identifying and recruiting for soft competencies. Despite the fact that these are not as simple to evidence, employers had devised a range of ways to assess these crucial attributes.

A number of these 'soft competencies' identified by our interviewees already align with data literacy skills, such as, *creative and analytical capacity*, *presentation*, and *problem solving*. Interestingly, one respondent said, *we prefer soft skills. Hard skills could be learned*, and it is interesting to view data literacy through this lens. In the same way, specific software can be easily learned, but broader critical skills are not so simply acquired.

Some areas in which organisations were already assessing crucial data literacy skills through soft competence assessments included a test for the ability to be a 'completer finisher' that reviewed the ability of candidates to understand what a document that was appropriate in detail for the chief executive officer might look like. Another looked for, *the ability to speak up... analysts can be on the introverted side so in a smaller company where we are very lean, so even if they are right or wrong I would expect an analyst to be ...able to voice their concerns.*

INSIGHT 5: While previous experience is prized, universities are seen as responsible for furnishing young employees with appropriate skills

Previous experience in a job and, *long and boring practice* are seen as desirable by employers, as this knowledge is perceived to be consolidated and internalised. It is acknowledged that younger and more junior employees, without experience, should be acquiring these skills at university, and both age (older employees will not have had the same opportunity to learn data literacy at university) and career stage (younger employees will not have had time to learn on the job) affect the expectations of employers.

At the same time, if employees are learning - or honing - their data skills on the job, this means it is harder for new skills to be acquired within industries. While some organisations or employers may make an effort to ensure new and innovative data skills are brought into the organisation, the cost of the time and financial resource this requires is prohibitive. Continuous Professional Development (CPD) was not suggested as a solution by many employers. This approach leaves a gap for new skills to be identified and acquired, and as more new employees enter the workforce, this gap can be filled by universities.

The data literacy gap

SME employers are relatively clear on what data literacy means to them when thinking about the concept in the abstract. However, when they discuss the data literacy skills they hire for, these become subsumed beneath role specific technical skills, which frequently include data. This leaves quite a substantial gap in the workplace for the softer, less role-specific, less frequently evidenced skills such as critical assessment of data, data collection and transforming analysis into insight, and in particular, the understanding of the ethical implications of the use of data.

The DEDALUS competence framework

Learnt in school/first job	**May be learnt through job experience**
Basic Digital Skills	**Basic Data Literacy**
Eg able to use Office, email, basic IT troubleshooting, proprietary or SaaS CRM or similar databases	Eg knowing what constitutes data within an organisation; basic management skills; simple interpretation of data and visualisations; basic compliance knowledge; ability to source data when required.
Position and role dependent, including SEO, data mining, analysis, technical aspects of privacy, security, visualisation - may not be associated with wider overview of data role in organisation or amongst stakeholders	Position and role dependent, including in depth knowledge of privacy and ethical considerations; visualisation theory, cleaning, curating and collecting data; critical assessment of data; source comparison; critical assessment of analysis; generating insights. Has overview of role of data in and outside organisation.
Advanced Data Skills	**Advanced Data Literacy**
Learnt in HE/through job experience	**May be learnt in HE/through job experience**

Table 2: Although both basic and advanced skills related to specific tasks and software are well taught, more general data literacy, at the advanced and particularly at the basic level, is not commonly provided

The data literacy gap in higher education in Europe

Table 2 shows that there is a distinct gap for the delivery of both quite simple and advanced data literacy programmes to support the needs of SMEs. At the basic level these will assist with data governance and management within the SME, as well as with communicating with external audiences such as customers and suppliers. At the more advanced level this will change the range of actionable insights that organisations are able to develop and execute.

We have identified three key actionable insights for data literacy programmes in HE to support SMEs in being able to access a suitably data literate workforce

- Ensure the ***provision of the full spectrum of data literacy***, and not just data skills;
    -
- Build on generic computer science modules with ***bespoke, relevant elements for specific fields and professions***, to help graduates develop the ability to extract value from data insights and application earlier in their careers;
    -
- Develop ways to ***evidence and certify the knowledge and insight*** aspects of data literacy, in order that employers are more able to identify and recruit for these softer skills.

In our next report we will review the HE and CPD training opportunity.

DEDALUS

Further reading (List to be completed once text is finalised).

[1] Millions of small businesses aren't GDPR compliant, our survey finds
https://gdpr.eu/2019-small-business-survey/

[2] Data Literacy: An essential skill for the industry LINK

[3] Full interview questionnaire (should put a copy on the website maybe?)

[4]Data Literacy: What is it, and how can we make it happen?
https://openjournals.uwaterloo.ca/index.php/JoCI/article/view/3274/4297

[5] Argast and Zvygintseva

[6] Woolf et al

ODI Data Literacy Report

(ADD MORE HERE)