data

MDPI

Review

Balancing Plurality and Educational Essence: Higher Education Between Data-Competent Professionals and Data Self-Empowered Citizens

Nils Hachmeister * , **Katharina Weiß, Juliane Theiß** and **Reinhold Decker**

BiCDaS, Bielefeld University, 33501 Bielefeld, Germany; katharina.weiss@uni-bielefeld.de (K.W.); juliane.theiss@uni-bielefeld.de (J.T.); rdecker@uni-bielefeld.de (R.D.)

* Correspondence: bicdas@uni-bielefeld.de or nils.hachmeister@uni-bielefeld.de; Tel.: +49-521-106-67599

Abstract: Data are increasingly important in central facets of modern life: academics, professions, and society at large. Educating aspiring minds to meet highest standards in these facets is the mandate of institutions of higher education. This, naturally, includes the preparation for excelling in today's data-driven world. In recent years, an intensive academic discussion has resulted in the distinction between two different modes of data related education: data science and data literacy education. As a large number of study programs and offers is emerging around the world, data literacy in higher education is a particular focus of this paper. These programs, despite sharing the same name, differ substantially in their educational content, i.e., a high plurality can be observed. This paper explores this plurality, comments on the role it might play and suggests ways it can be dealt with by maintaining a high degree of adaptiveness and plurality while simultaneously establishing a consistent educational "essence". It identifies a skill set, data self-empowerment, as a potential part of this essence. Data science and literacy education are still experiencing changeability in their emergence as fields of study, while additionally being stirred up by rapid developments, bringing about a need for flexibility and dialectic.

---

1. Introduction

Data are of the utmost importance throughout most facets of life, from academia [1–6], to politics [7,8], to the economy [9,10], to our daily lives [11]: it is generated by our cars, heating, fitness devices and communication. Its importance is still increasing as the costs of data generation are continuously dropping [12] and data covers more and more facets of the modern world. Equally important, sharing data, i.e., copying and transferring data, has become effortless and low-cost, even becoming cheaper at an exponential rate [13]. Furthermore, our societies are increasingly demanding data. Any claim, reasoning, decision or political measure is perceived to be more convincing if it is grounded in data.[1] This notion is exemplified by W. Edwards Deming's quote: "In God we trust, all others must bring data." Or, as Koltay et al. put it, "There is an aura of truth, objectivity, and accuracy around it [...]" [14].

Koltay and colleagues later on warn against falling prey to overly optimistic (or pessimistic) expectations regarding data. Objectivity, truth and accuracy (or their opposites) are attributes to be applied to a given analysis of data, not to the data itself. Notwithstanding, errors can be made already in data collection, which hampers the expressiveness of any analysis conducted on such data.

That is not to say that the increasing importance of data is unjustified. The "datafication" of our world is a powerful means of controlling, maintaining and improving that

---

[1] Or credibly appears to be grounded in data.

same world. However, we can only make the most of data when its potential and limitations and correct handling, analysis and interpretation are known. Known in every detail by a few experts excelling in data handling, but also known, on a less detailed level, by the general public. The importance data can have in public debate can be observed in many countries in Europe and elsewhere, in particular in the collective actions taken to counter SARS-CoV-2 [15,16]. This debate was centred around different statistical indicators, e.g., newly infected, deaths, hospitalisation and intensive care. At the beginning of the pandemic, discussion in public media was often on raw, absolute numbers of limited expressiveness, but then matured to more useful figures such as reproduction rates (R-value), normalised numbers and the statistical expressiveness of such figures became a subject of debate, examples for sources which attracted larger public interest in Germany and Europe are an interactive coronavirus map by the weekly newspaper *Die Zeit* [17], the Podcast Coronavirus-Update [18] or the COVID-19 Dashboard by CSSE, Johns Hopkins University [19]. The role data plays in fighting this pandemic, self-evident to any epidemiologist, has increasingly become a matter of common sense in public debates as well and is comprehensively summarised in Letouzé et al. [20]. Of course, data are being used for good and for ill during this debate, as many ill-formed statistics and data visualisations (intentional or not) were seen [21–23].

Accordingly, our educational systems need to adapt to these challenges, today more than ever. Curricula need to be developed, revised and adapted, modes of teaching data competencies need to be established, evaluated and enhanced. While this paper focuses on academic/higher education, it is the conviction of the authors that the teaching of data competencies needs to start much earlier, i.e., at the school level, in an adequately audience-tailored manner [24,25]. Therefore, teachers constitute an important group to be addressed by data literacy programmes. Data literacy programmes and publications that already address teachers—specifically or among other groups—focus on two different main areas namely (1) preparing teachers for teaching data competencies to students directly, e.g., [26,27], Bowen and Bartley (2014), as cited in [28], or (2) enabling teachers to improve their teaching by data-driven decisions [29,30]. Thus, teachers can act, directly or indirectly, as multipliers for data competencies and, even more importantly, for a general data awareness and data culture, directly by teaching data competencies to students within and across different disciplines and indirectly by data-driven guidance of teaching.

Furthermore, it is important to bear in mind that due to differences in educational systems, curricula development for data competencies differs around the world. In consequence, academic discourses are at least partly separated, as for instance between European countries and the US. Due to the authors' backgrounds, in the present paper, we will put an emphasis on the discussion in Europe, especially in Germany, while also referencing discourses in other countries, especially the US and Canada, and pointing out some noteworthy differences.

Academic education differentiates between disciplines and degrees (it is audience-tailored, if you will). It is governed by (an abstract understanding of) the requirements of the professional worlds students will eventually join. For the same reasons and by the same factors, education in data competencies needs to be differentiated. Obviously, a historian will need different data competencies[2] than an astrophysicist. Teachers, as multipliers to the next generation, have received focus in this debate, particularly in the US [29–31]. In the area of teaching and leadership in education, four different roles for data experts that highlight different aspects of data expertise in the context of the US educational system have been defined: (1) practitioner administrator, (2) educational quantitative analyst, (3) research specialist and (4) data scientist [32]. Each of these four types would need its own, tailored set of data competencies. In Germany and other European countries, where the educational system is much less driven by standardised assessments and their results (that is data), these roles do not have a one-to-one correspondence.

---

[2] And it has only recently been recognised in the first place by a wider community that data competencies might come in handy to historians.

The distinction between data science and data literacy education has been broadly discussed in recent years. Both are active topics of academic discussions and curricula development around the world (e.g., in Germany [15,33,34], in Canada [35] or the US [29,30]). The distinction between the two is often clarified by the analogy of elite sports (≈data science education) and mass sports (≈data literacy education).[3] Like every analogy, this has some shortcomings; for example, mass sports rarely has a professional dimension, while for data literacy a fundamental assumption is that it is needed in the professional world. It also implies seeing data literacy as a 'minor data science', a notion we do not share, as we will discuss on several occasions throughout this paper. It also has its strengths, such as in implying that neither data literacy nor data science education ends after receiving a degree.

Metaphors aside, describing the difference between data literacy and data science education can be done from a variety of perspectives. Considering the field of tension between a given domain and its data (handling), the difference between the two is certainly in where on that field the focus lies [29,30,36]. Data literacy education is the interdisciplinary cornerstone for critical and autonomous data handling during the whole process of transforming data into (actionable) knowledge in a certain domain, in which models and methods are means to an end. On the other end of the spectrum, data handling, methods and models are central to a data scientist and are being taught (or should be taught) on a sufficiently general level that they are able to apply those skills to varying domains [4]. As a result, data scientists also need to be well-versed in audience-tailored communication with professionals from other fields [9,38].

The level of expertise in a given data competency, such as data collection and preparation, forms a continuum from novice to expert. This idea of a continuum is a reoccurring theme in the literature on data literacy education [39,40] and has been similarly proposed for other literacies as well [41]. Virtually, any degree of competence can be found in some individuals for a given data competency and the degree of competence needed by different data literate professionals varies across domains. To describe these different needs with respect to this (not quantifiable) continuum, authors like Beck at al [42] have introduced discrete levels of expertise, verbally locating them on the continuum of competence. Generally speaking, data scientists will be found more towards the expert end of that continuum for most data analysis methods and models. This is not necessarily true for other data competencies, as we will discuss towards the end of Section 2. Additionally, for rather narrowly defined competencies with a high demand for some discipline (e.g., time series analysis in physics) it is possible that a data literate professional of that domain could outperform any data scientist[5].

As such, data-competent domain experts are key in the transition of data into knowledge and decisions. Thus, the knowledge, skills—across the data life cycle—and mindset that are required for managing and understanding data and data products—referred to as data literacy—have become increasingly important for graduates from all backgrounds [15,33,35], particularly teachers [29,30] (as already mentioned above). Furthermore, there is a wide range of literature on data literacy and closely related topics such as data information literacy [43] targeting different audiences within and beyond academia. Target audiences within academia in addition to researchers or educators include graduate students enrolled in research courses or interested in the experimental method [44,45] or academic librarians [43]. Beyond academia, target audiences include the general public [16,46], teachers, school administrators and high school librarians [29,47,48] and professionals [49].

---

3 This analogy has, for instance, been a reoccurring theme at the Data Literacy Education Symposium by Stifterverband and GI in 2018 (https://www.stifterverband.org/veranstaltungen/2018_04_24_data_literacy_education_symposium).

4 "The best thing about being a statistician is that you get to play in everyone's backyard." John Tukey [37], before the term data science was invented, but it applies to data science just the same.

5 This is one point in which regarding data literacy as a minor data science would be inadequate.

All over the world, universities [44], schools [45,48] and the private sector [49] have recognised the requirement to train students and employees in domain-specific data skills. The aim is to empower them to autonomously and critically handle data and data products in their specific field of expertise [2,3,5,6,31,50,51]. This need to empower citizens and professionals with domain-specific data skills has also been recognised in countries of the global south, e.g., Mexico, Colombia and Brazil [52].

In this paper, we provide a commentary and discussion on different aspects of data-centred higher education—both data literacy and data science. We observe a plurality of educational contents and will comment on the role this plurality might play, for good or ill. We will make a point of balancing plurality of educational content with the need for an educational essence and will suggest a set of skills we summarise as data self-empowerment (structured along the entire data life cycle) to constitute (an important part of) the essence of data literacy education.

Related to the educational plurality we observe is the active debate on definitions of central terms of the field, particularly the term data science itself, and the rapid technological development in this field [4]. Both will compel academic, data-centred education to continually adapt and update its educational content, more than many other disciplines.

2. Goals of Higher Education Regarding Data Competencies

What are the goals of academic data science and data literacy education? First and foremost, of course, these are the goals of academic education in general. Universities have a societal mandate to educate their students and this mandate is, from our perspective, three-fold: (i) equipping students with the skills required for the professional world in their chosen domain, (ii) equipping students with a basic scholarly mindset and the tools to eventually start an academic career, and (iii) preparing students to fulfil roles as active, responsible, thoughtful members of society. [53].

Equipping students with the data skills required for the professional world (i) and a basic scholarly mindset (ii) requires different (sub-)sets of data competencies to be taught, depending on the discipline (data competencies required of a historian $\neq$ those required of an astrophysicist, as noted in Section 1). For a data scientist, this set of competencies is that of a data generalist, education regarding methods and models being more in-depth, considering diverse applications and also training in communication with professionals from various domains.

The goal of educating active, responsible members of society (iii) is different insofar as the required competencies target life beyond the professional and academic world, and thus the same skills can be taught mostly independently of the main subject of study.[6] Beyond the professional and academic world, we identify (with respect to data) three main entities which play a role: (i) the individuals themselves[7] and the social groups they consider themselves to be a part of in some given context, (ii) society at large, within which (public) debates are being conducted that culminate in the media, some of which might affect/concern the individual or his/her social groups. In these debates, data are increasingly being used on all sides, sometimes with competence and thought, sometimes in ignorance and sometimes with ill intent (see below). In recent times, data, the use of data and (unintended) consequences of the use of data products have been the subject of public debate themselves. The term 'society' here also encompasses politics at different levels, which uses data and which can be held accountable using data [39]. Some players might have interests opposing the individual's interests (that of his/her social groups), who therefore become (iii) adversaries of the individual. This encompasses persons and organisations performing unwanted information gathering about the individual (tracking),

---

[6] However, the mode of teaching (pace, level of detail, etc.) may be adapted based on educational background, i.e., main subject of study. Also potential synergy effects with the main subject should be realised, of course.

[7] The persons whose data skills are under consideration/students being educated.

seeking illegal computer access to the disadvantage of the individual (hacking) and those seeking to disinform the individual in pursuit of some (political) agenda ('fake news').

Navigating this field in an increasingly data-driven world requires a dedicated set of skills; not falling prey to misinformation requires solid skills in the (critical) reception of data analysis. Not giving away some information unintentionally requires knowledge and skills about tracking and effective counter-measures as well as a basic understanding of applicable data protection laws with reference to one's own rights and powers. The aim is not to keep personal information secret at any cost, but rather to consciously decide which information to disclose and for what (personal) benefit [46]. To not fall prey to hacking, basic skills in digital self-defence are needed. Partaking in societal discussions relevant for the individual requires basic skills in the interpretation and critical reception of data analyses [54] manipulation of data and the creation of data products. And, as data and data products are increasingly often the subject of societal discussion itself, a (substantiated) anticipation of consequences of the use of data (data ethics) is key. A foundation for all of this is a general awareness of the importance of data, of challenges and questions arising around it [39,46].

We, therefore, define the educational goal data self-empowerment that aims to equip students with all data-related skills and knowledge required for:

a   Partaking in societal discussions by employing data to substantiate arguments and counter others' arguments.

b   Developing educated opinions when the use of data and data products is the subject of societal discourses itself.

c   Defending oneself from others who aim to use data and information technology to the disadvantage of the individual.

d   Critically assessing claims of others, particularly when these claims appear to be grounded in data i.e., countering 'fake news'.

e   Assessing political decisions based on data and hold accountable government officials.

These skills are highly relevant today and will become increasingly crucial for societal participation in the foreseeable future [4].

In the panel discussion concluding the 2020 German Data Science Days at LMU, Munich, K. Schüller pointed out that it is anything but self-evident that today's data science graduates have acquired the skills we summarise here under the term data self-empowerment.

While fully admitting that other notions are possible, these arguments, combined with the aforementioned domain-independence of data self-empowerment, compel us to see data self-empowerment as an essential part of data literacy education. Consequentially, data scientists in training, indeed, need to undergo an (audience-tailored) data literacy training, just as students in history, physics, medicine or any other domain.

3. Exploration of the Content of Data-Related Education

During an exploratory analysis in early 2020, we had a closer look at a variety of data science study programs. In doing so, we noticed a huge plurality in existing educational content.

While being applied to many different fields, data science draws on three main fields regarding methods and models: statistics [55,56], machine learning [57,58][8] and mathematical modelling [36]. Different study programs set different focal points within the field spanned by the three, causing at least in part the observed plurality in the educational content of different data science study programs.

To illustrate this plurality (which can equally be observed at the Master's level), in Table 1 the educational content of courses of four Bachelor level data science study

---

[8]   Actually Donoho [58] nicely portrays the distortions the debate over the interrelation between statistics, (machine learning) and data science has seen.

programs at different universities have been listed as an example. These are highly focused on statistics, mathematical modelling or machine learning and nicely illustrate that, while there is certainly consensus about the general goals of data science education, the interpretation of these goals (and of the term data science) can be disparate.

This disparity has many reasons, one of which can be observed in the second line of Table 1: while actually being quite a trans-disciplinary topic, data science study programs are often hosted by one academic department or, in rare cases, an alliance of two or three departments. It is certainly only natural that any group of lecturers tasked with devising a new study program will introduce their own background and expertise on many different levels.

Table 1. Educational contents (translated to English) of exemplary data science Bachelor courses, illustrating the differences in educational contents. For TU Dortmund [59] we see a clear focus on statistical methods, the degree from Saarland University [60] focuses much more on programming and AI, while in the program from the University of Warwick [61] many mathematical topics can be found, complemented by programming and statistics. The study program of the highly prestigious UC Berkeley [62], who also is a precursor in data literacy education, incorporates a domain focus, following the widely accepted notion that data science studied detached from any domain is not data science.

TU Dortmund, D	Saarland University, D	Warwick University, UK	UC Berkeley, US
Dep. of Statistics Dep. of Computer Science Dep. of Mathematics	Dep. of Mathematics and Computer Science	Dep. of Statistics	Div. of Computing, Data Science, and Society
Mathematics I & II Data structures and algorithms I & II Introduction to statistics and data science Probability theory Programming Estimating and testing Scientific work Introduction to statistical learning Statistical methods Software applications Management of large data sets Project work	Programming I & II Mathematics for computer scientists I, II, III Lecture series Elements of data science and artificial intelligence Statistics lab Applications I & II Elements of machine learning Basics of theoretical computer science Basics of data structures and algorithms Big data engineering Project seminar data science and AI Core lecture data science and AI I & II Advanced lecture data science and AI I & II Seminar data science and AI Bachelor seminar	Programming for computer scientists Design of information structures Mathematical programming I Linear algebra Mathematical analysis Sets and numbers Statistical laboratory 1 Introduction to probability Mathematical techniques Database systems Algorithms Software engineering Stochastic processes Mathematical methods Mathematical statistics part A & B Data science project	Foundations of data science Calculus I Calculus II Linear algebra Program structures Data structures Domain emphasis Data 100: principles and techniques of data science Computational & inferential depth Probability Modeling, learning, and decision-making Human contexts and ethics Domain emphasis

Note: US study programs are not easily comparable to those of European universities for different reasons, e.g., significant differences in the educational systems.

In many countries, especially in Germany, data literacy curricula in higher education are only just starting to develop, and data literacy competencies are being integrated into existing discipline-specific courses or taught in novel interdisciplinary programmes for students at different levels. Surveying the different data literacy education programs offered, a large plurality can easily be spotted. We have to differentiate between three different sources of plurality, though: (a) intended or necessary plurality due to audience-tailoring the teaching content to the main subject, (b) necessary plurality due to differences in educational systems of different countries as well as their cultural and economic background, applicable (data protection) law, political agenda and other aspects and (c) unintended (possibly undesirable) plurality caused by different approaches in curricular design. The last of the three might blur the term of a data literate professional, and potentially pose a hurdle to prospective staffing processes (compare Section 6).

4. The Need for a Clear Definition

In fact, there is currently no widely accepted definition of data science [54,57,58,63] and the above-mentioned plurality in educational content is certainly (at least partially) rooted in this ambiguity. Several divergent definitions can be found, e.g., in [9,12,36,55,64]. Occasionally, authors define the term data scientist instead, as in [9,38].A strong, widely accepted definition of the term data science would certainly form a foundation for coherent, well-grounded data science teaching, but lecturers need to make decisions on the teaching content in courses offered now. In the current fast paced times, there is little alternative: businesses and society are demanding data scientists now [9,38,58,65,66]. Waiting for the long-lasting, comprehensive discussion on a definition for data science to be settled is not an option.[9] It seems strange, though, to teach a newly emerging discipline (data science) before there is even a widely accepted consensus on what constitutes this discipline[10] .

There are, thus, two parallel, active discussions being led: (a) on the definition of the term data science and (b) on the teaching contents a thorough data science education needs. The latter of these discussions is certainly more directly connected with our topic here and we will give an overview about several seminal contributions, which received much attention in the European/German data community in Section 6.

There is a similar discussion about the teaching contents of data literacy education. However, this discussion is naturally structured, as is data literacy in higher education, into general data competencies, particularly those required for data self-empowerment and domain-specific competencies.

Furthermore, the discussion about teaching contents seems also influenced by noticeable differences in educational systems. In its extremes, this can lead to topics being avidly discussed in one part of the world while hardly being of any concern in some other: in the US, where standardised students assessments are far more common than in Germany, the question whether assessment literacy should be seen as part of data literacy is broadly discussed, whereas assessment literacy plays a rather minor role in the German discussion [29,30,67].

Data literacy is often defined as the ability to collect, manage, evaluate, and apply data in a critical manner [35]. Although this definition is widely accepted, discussion(s) about which competencies are required to master these abilities and about their prioritisation are ongoing [15,29,30,33,35,40,68]. For recent overviews of different competence frameworks, see [53,67]. There is a related active discussion about the boundaries of data literacy education with respect to other literacies such as digital literacy, information literacy or statistical literacy, e.g., [14,69], or more strongly discussed in the US—, with respect to assessment literacy, e.g., [29,30,67]. For a recent extensive discussion of different literacies see [70].

The need to determine teaching contents for each domain is one central driver for plurality of data literacy teaching content. While in data-focused domains, such as physics or economics, the collection, management, evaluation and application of data already make up an integral part of the curricula, these abilities only play a minor role in the curricula of less data-centred disciplines, such as history or law. Nevertheless, even in such traditionally less data-centred disciplines, data and data analysis have become increasingly important as well as commonplace. This is illustrated by the emergence of the field of digital humanities as well as by new challenges posed for these disciplines by datafication and technological innovations. For instance, in law, new data-related challenges are provided by autonomous systems and artificial intelligence [71]. Faculties and universities are adopting this development as well, as exemplified by the emergence of the Institute

---

[9] Even more so, as technical advances in general and the fast-paced development of the field of data science in particular will prevent this discussion from converging any time soon.

[10] What future implications this 'premature education' has, remains to be seen. We are either currently cementing the current ambiguity of the term data science or, within a few years, there might be those who have studied data science, but whose skill set no longer matches the contemporary definition of the field.

for the Law of Smart Systems at Bielefeld University.[11] This also changes the publication landscape of the discipline, as is exemplified by the established and prestigious publisher Taylor & Francis, who, in 2009, dedicated a new journal to these arising challenges: *Law, Innovation & Technology*.

However, despite the described plurality in required skills the idea of data literacy is not a cornucopia of skills to pick from, but should be regarded as the foundation for domain-specific data handling, critical thinking, problem solving and interdisciplinary collaboration. Thus, the definition of data literacy should not stop at the level of skills but should also encompass the knowledge, abilities, motivation and mindset that are required for the autonomous handling of data in different domains.[12]

The current status quo thus presents itself as a strong plurality in teaching contents rooted in a lack of a common understanding of what data science or data literacy are.

5. Plurality in Data Science and Data Literacy Education

What impact has this plurality on data science teaching? First things first: neither extreme of plurality and coherence are desirable. Too much plurality diverges to arbitrariness of educational content, making the qualification of a data scientist/data literate meaningless in a professional context. On the other hand, completely coherent educational contents would deprive students of the possibility to pick focus areas or to develop a professional profile and a distinct skill portfolio. Additionally, the field of data science is currently undergoing a highly dynamic development, requiring some degree of flexibility in, and continuous reevaluation of, educational contents/curricula.

Thus, data science education needs to develop an 'educational essence'. This is a set of soft and hard skills and knowledge that each and every data scientist should have. The skills and knowledge encompassed by the educational essence of data science degrees should remain relevant irrespective of the future technological development of the field, within the limits of the human ability to forecast said developments. This educational essence creates reliability for prospective employers. It also helps to define the discipline of data science as a whole. Discussion on what this educational essence might be and along which lines students might develop professional profiles is active and ongoing, and we will look at some major contributions in Section 6.

On the other hand, a huge portion of the current plurality ought to be maintained, leaving the flexibility to differentiate and build an individual profile, appealing to certain (groups of) potential employers. Balancing a rigorous essence of data science education against a high plurality is a challenge. To make things more involved, due to the ongoing debate about the definition of data science, as well as the fast-paced technical developments in the field, this balance is continually disrupted and needs to be re-evaluated, resulting in adaptations of curricula and skills being taught.

For the young and dynamic field of data literacy education, a plurality in educational content between different countries, between different institutions of higher education and even individual actors can be observed, but the concept of an educational essence can serve data literacy just as well. As for data science education, this plurality provides opportunities, e.g., in meeting the rapid advancements in the field and in individual profile development or in the plurality helping to complete the exploration stage characterising the state of data literacy education, at least in some countries and contexts. But it also holds challenges, such as in establishing a reliable certification of data literacy competencies. Plurality in data literacy education is inevitable on the domain-specific level. For example, the usage of data analysis tools for qualitative versus quantitative data or the knowledge about data sources differs significantly between the various academic fields. However, the educational essence of data literacy is not the expert application of specific methods, but rather comprises the knowledge, skills and mindset to gain information from raw data

---

[11] See http://www.jura.uni-bielefeld.de/institute/rit/was_rit.

[12] This is another point at which we feel it would be wrong to regard data literacy as the 'minor data science'.

or data products and to responsibly translate this knowledge into action. Thus, maintaining a certain degree of plurality in data literacy education is undoubtedly desirable, but care should be taken to prevent the dilution of the concept of data literacy, especially in the advent of ever newly arising 'literacies' [50]. In Germany, a national funding program promotes the development of data literacy initiatives at higher education institutions which are considered to be lighthouse projects for data literacy education in Germany, but in their plurality are also fields of experimentation. The actors from these different data literacy programs join forces in a nationwide data literacy network by sharing ideas, concepts and experiences. In this environment, Bielefeld University, together with the University of Paderborn and the University of Applied Sciences in Bielefeld, has developed a concept to guide students and teachers in Germany through the 'jungle' of data literacy competencies and to provide graduates with a comprehensible evidence of their data skills, a data literacy certificate is currently being developed within the DaLiS@OWL-project[13] that frames the plurality to allow individual and domain-specific specialisation without diluting the educational essence of data literacy education.

6. Competence Frameworks—An Active Discussion from a European Perspective

In both fields, data science and data literacy education, there is a variety of contributors to the debate about educational contents. We will briefly introduce some (without claiming or aiming at completeness), which we consider particularly seminal from a European and German perspective and/or have been much received in the German and European communities. For data literacy, the following publications are certainly noteworthy:

Strategies and Best Practices for Data Literacy Education—Knowledge Synthesis Report: In 2015, Chantel Ridsdale and colleagues from Dalhousie University published a comprehensive and much cited report about data literacy education compiled from various sources ranging from peer-reviewed publications to governmental reports, grey literature and informal blogs [35]. From a systematic review of the existing literature, the authors synthesise a widely accepted definition of data literacy as "the ability to collect, manage, evaluate, and apply data, in a critical manner". Furthermore, this report presents a data literacy competencies matrix organised by the five elements of the data literacy definition. For each of the elements of their definition, the authors describe competencies, skills, knowledge and expected tasks which are categorised into conceptual, core and advanced competencies. The report also provides best practices for teaching data literacy education at the university level as well as an extensive annotated bibliography on data literacy. Overall, the Knowledge Synthesis Report can be considered as the groundwork for many data literacy initiatives and a major impulse for current developments in this field.

Future Skills: Ein Framework für Data Literacy: (~'*Future Skills: A Framework for Data Literacy*') German publication by Schüller and colleagues [33] introduces a novel data literacy competence framework which adds a new perspective to the data literacy competencies matrix of the Knowledge Synthesis Report. Prior to the development of their competence framework, Schüller and colleagues carried out a detailed review of the available literature, supplemented by systematic interviews with experts. Besides definitions and competencies, the authors also focused on testing instruments for data literacy. The methodology and the findings of the systematic review were published separately [53]. On this basis, a novel data literacy framework was developed which distinguishes itself not only by describing data literacy competencies, but also by integrating them in a cyclic process of producing and receiving steps. In this model, data literacy competencies, skills and mindset are described on a 'coding level' (for the generation of data products) as well as on a 'de-coding level' (for the interpretation of data products). An updated English version of this framework [15] was recently published which includes a discussion of the need for data literacy education during

---

13   For more information see https://www.uni-bielefeld.de/datascience/teaching/data_literacy.html.

the current SARS-CoV-2 pandemic as well as opportunities the pandemic could provide for furthering data literacy education. Thus, although the publications presented so far describe similar competencies for data literacy, they differ considerably in the categorisation of these competencies with significant implications on methods for teaching and testing.

Data Literacy for Teachers: In their seminal framework, Ellen Mandinach and Edith Gummer address data literacy education for a specific group of professionals: teachers or educators [29,30]. It incorporates skills, knowledge and dispositions that are specifically relevant for making data-driven decisions in the context of classrooms. This is already reflected in the authors' definition of data literacy which highlights types of data that are especially relevant in school contexts and forms of knowledge that are important for teaching children: "Data literacy for teachers is the ability to transform information into actionable instructional knowledge and practices by collecting, analysing and interpreting all types of data (assessment, school climate, behavioural, snapshot, longitudinal, moment-to-moment etc.) to help to determine instructional steps. It combines an understanding of data, disciplinary knowledge and practices, curricular knowledge, pedagogical content knowledge, and an understanding how children learn" [29] (p. 14). Firstly, the complex framework incorporates seven knowledge components: (1) content knowledge, (2) general pedagogical knowledge, (3) curriculum knowledge, (4) pedagogical content knowledge, (5) knowledge of learners and their characteristics, (6) knowledge of educational contexts, and (7) finally knowledge of educational ends, purposes and values. These knowledge components interact with the domain 'data use for teaching' which constitutes an iterative inquiry cycle with five sub-components (1) identify problems and frame questions, (2) use data, (3) transform data into information, (4) transform information into a decision, and (5) evaluate outcomes. Underlying these five sub-components of 'data use for teaching' is a large set of more than 50 specific skills, knowledge (and dispositions) needed by educators. Although the framework was developed in the US and some components are especially relevant in this educational system, it is also relevant for the development of teacher training curricula in general.

The above-described publications provide detailed analyses of ongoing developments in the field of data literacy education and well-grounded competence frameworks. Furthermore, they provide best-practice examples for educating students to critically and autonomously act in the world of data on a professional level as well as responsible citizens. However, the implementation of data literacy content in the curricula as well as the development of discipline-specific and interdisciplinary courses still pose significant challenges for teachers and universities as organisational structures. Awareness of data literacy in higher education is increasing in Germany and other countries, as it can be observed on the level of public funding for universities, a growing number of commercial data literacy courses as well as in the increasingly often stated requirement of data literacy skills across all business sectors.

Different approaches have also been made to define what data science curricula need to contain and the results differ depending on the approach taken and also the scientific field/domain of the authors. We will illustrate this by the example of three seminal contributions of the recent past.

Data Science: Lern-und Ausbildungsinhalte: (∼*Data Science: Didactic and Educational Contents*)—In 2018 the Gesellschaft für Informatik (∼*German Informatics Society*) has founded a Task-Force 'Data Science/Data Literacy', which has since yielded several publications on the topic. The most recent publication [36], from December 2019, identifies fourteen different categories of relevant competencies. Applying an Anderson-Krathwohl-Taxonomy [72], three ideal-typical personas typically found among prospective data science Master's students are described (further refined to five personas in a second step): (i) a student who has finished a Bachelor's degree in statistics, mathematics or similar, aiming to complete a Master's in data science and to

later work in industry or research, (ii) a student who has finished a degree in a domain field (science/humanities) aiming to acquire strong analytic skills to apply in that domain, (iii) a professional of some domain with years of work experience, aiming at a hinge function between domain professionals and data scientists. Orthogonal to this, three levels of expertise have been defined:[14] understanding, application, analysis. The publication contains a comparably explicit discussion of educational goals, structured via the different personas. It takes into view different study programs of different German universities and universities of applied sciences and exemplifies the needs of the different personas in terms of the different foci of the study programs. It also offers a good overview of different relevant publications. Compared to other competence frameworks this is rather lightweight, yet employs a highly systematic approach and is recognisably written from an informatics perspective.

EDISON: Funded by the EU, the EDISON project has developed what is possibly the most elaborate approach. The project has produced several interdependent documents, of which the competence framework data science [73] constitutes a cornerstone. It references several competence frameworks of adjacent fields and tries to establish compliance with these. It identifies a total of five fields of competencies (including sub-divisions of these fields) and, based on that, defines a series of skill sets taking different perspectives. It also offers guidance in how to apply the framework. Another central document is the body of knowledge [74], listing six groups of relevant areas of data science knowledge. Again, each area then contains different knowledge units. The EDISON approach is detailed and elaborate, which also means that it requires quite a substantial amount of work to digest.

Vermittlung von Datenkompetenzen an den Hochschulen: Studienangebote im Bereich Data Science: (∼*Teaching Data Competencies in Higher Education: educational Offers in the Field of Data Science*)—A study by the (German) HIS-HE from 2018 [75] assesses the current needs in the field of data science professions and compares it with existing data science study programs, deriving recommendations for future developments in that field. The main focus of this publication lies in satisfying the demands of the job market. The authors analysed job postings as well as information about data science relevant study programs in Germany and conducted stakeholder interviews. In the end, seven recommendations were derived addressing different stakeholders, particularly higher education and enterprises – from forming networks to summer schools for diffusing data science skills beyond the core disciplines.

These three publications tackle the topic of data science education from three different perspectives, namely from the employers' perspective, from a methodological perspective and from the employees' perspective. They also diverge in the methods used: analytic thinking, surveys/interviews and qualitative analysis, text analysis and clustering. While all three are noteworthy publications, none can definitely conclude the topic.

7. Conclusions

The status quo in higher education is currently characterised by a vast plurality of educational content subsumed under the terms data science and data literacy education. This holds true for German higher education and, in varying degrees, for many other countries. We identify some main roots of this plurality in (i) the diversity of faculty homes of these study programs, (ii) the rapid technological and methodological development, (iii) regional, national, legal, economic, educational and cultural differences in educational systems and (iv) the lack of a widely accepted definition of the concept of data science in data science education. Data literacy in particular is more susceptible to regional/national differences in educational systems, because primary education is much more nationalised/regionalised as compared to higher education.

---

[14] Again landmarking a continuum of expertise, compare Section 1.

This plurality can be beneficial, given the need for students to build unique and attractive skill portfolios. In order to create reliability for prospective employers and to not dilute the terms of data literacy and data science, a common educational essence should be developed for each of these fields. Different contributors continue to pursue that goal by publishing competence frameworks and bodies of knowledge.

We support the notion that data literacy is a qualification in its own rights, rather than a minor/reduced version of data science. Consequently, we believe that data science students can benefit from data literacy education. Particularly, the skills required for data self-empowerment, i.e., skills that empower the individual within a data-driven society and against attempts to use data to the individual's disadvantage, are typically not part of today's data science curricula yet might characterise the educational essence of data literacy.

Circumstances are compelling us to already educate students in the field of data while this field is stirred up by rapid technological and methodological developments and a whole new field of science is emerging. This raises a variety of challenges: ambiguity about definitions, plurality of educational content along several dimensions and a need to periodically update curricula, to name a few. Addressing these challenges requires agility and institutional efforts, involving data practitioners and lecturers from within and beyond the university. We believe that recollecting on the societal mandate of higher education can provide a fixed landmark in this field in turmoil and in contextualising this mandate in the field of data, we propose data self-empowerment as an educational goal beneficial for students, at schools as well as at higher education institutions, across disciplines.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Gibson, J.P.; Mourad, T. The growing importance of data literacy in life science education. *Am. J. Bot.* **2018**, *105*, 1953–1956. [CrossRef] [PubMed]
2. Manovich, L. Data science and digital art history. *Int. J. Digit. Art Hist.* **2015**, *1*, 12 – 35. [CrossRef]
3. Molina, M.; Garip, F. Machine learning for sociology. *Annu. Rev. Sociol.* **2019**, *45*, 27–45. [CrossRef]
4. Grillenberger, A.; Romeike, R. Developing a theoretically founded data literacy competency model. In Proceedings of the 13th Workshop in Primary and Secondary Computing Education, Potsdam, Germany, 4–6 October 2018; Association for Computing Machinery: New York, NY, USA, 2018; pp. 1–10. [CrossRef]
5. Kalidindi, S.R.; De Graef, M. Materials data science: Current status and future outlook. *Annu. Rev. Mater. Res.* **2015**, *45*, 171–193. [CrossRef]
6. Ruehle, F. Data science applications to string theory. *Phys. Rep.* **2020**, *839*, 1–117. [CrossRef]
7. Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions 'A European Strategy for Data', 2020. Available online: https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1593073685620&uri=CELEX%3A52020DC0066 (accessed on 17 September 2020).
8. Eckpunkte Einer Datenstrategie der Bundesregierung [Cornerstones of a Data Strategy of the Federal Government of Germany], 2019. Available online: https://www.bundesregierung.de/resource/blob/975226/1693626/60b196d5861f71cdefb9e254f5382a62/2019-11-18-pdf-datenstrategie-data.pdf?download=1 (accessed on 17 September 2020).
9. Stockinger, K.; Stadelmann, T.; Ruckstuhl, A. Data Scientist als Beruf [Data scientist as a profession]. In *Big Data: Grundlagen, Systeme und Nutzungspotenziale [Big Data: Foundations, Systems and Utilisation Potentials]*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 59–81._4. [CrossRef]
10. Nagel, C.; Litzel, N. Data Scientists—Heiß Begehrt Auf Dem Arbeitsmarkt! [Data Scientists—Highly Demanded on the Job Market!]. Bigdata Insider. 2020. Available online: https://www.bigdata-insider.de/data-scientists-heiss-begehrt-auf-dem-arbeitsmarkt-a-708584/ (accessed on 17 September 2020).
11. Bowler, L.; Acker, A.; Jeng, W.; Chi, Y. "It lives all around us": Aspects of data literacy in teen's lives. *Proc. Assoc. Inf. Sci. Technol.* **2017**, *54*, 27–35. [CrossRef]
12. Van der Aalst, W. Data Science in action. In *Process Mining: Data Science in Action*; Springer: Berlin/Heidelberg, Germany, 2016; pp. 3–23._1. [CrossRef]

13. Aslan, J.; Mayers, K.; Koomey, J.G.; France, C. Electricity intensity of internet data transmission: Untangling the estimates. *J. Ind. Ecol.* **2018**, *22*, 785–798. [CrossRef]

14. Koltay, T. Data literacy: In search of a name and identity. *J. Doc.* **2015**, *71*, 401–415. [CrossRef]

15. Schüller, K. Future Skills: A Framework for Data Literacy—Competence Framework and Research Report. *Zenodo* **2020**. [CrossRef]

16. Jones, B. *Data Literacy Fundamentals: Understanding the Power & Value of Data*; Data Literacy: Bellevue, WA, USA, 2020.

17. Blickle, P.; Dinklage, F.; Engmann, R.; Erdmann, E.; Fischer, L.; Gortana, F.; Klack, M.; Kreienbrink, M.; Loos, A.; Peter, V.; et al. Coronavirus: Welche Regionen Besonders Betroffen Sind [Corona Virus: Which Regions are Particularly Affected]. Die Zeit [The Times]. 2020. Available online: https://www.zeit.de/wissen/gesundheit/coronavirus-echtzeit-karte-deutschland-landkreise-infektionen-ausbreitung (accessed on 17 September 2020).

18. Hennig, K.; Martini, A.; Drosten, C. Coronavirus-Update. Podcast, 2020. Available online: https://www.ndr.de/nachrichten/info/podcast4684.html (accessed on 17 September 2020).

19. COVID-19 Dashboard by the Center for Systems Science and Engineering (CSSE) at Johns Hopkins University (JHU). Website, 2020. Available online: https://coronavirus.jhu.edu/map.html (accessed on 17 September 2020).

20. Letouzé, E.; Oliver, A.; Bravo, M.A.; Shoup, N. Using Data to Fight COVID-19 And Build Back Better. Technical Report, Data POP-Alliance and Vodaphone Institute for Society and Communications, 2020. Available online: https://www.vodafone-institut.de/wp-content/uploads/2020/10/VFI-DPA_Fighting_COVID_with_Data_report_2020.pdf (accessed on 17 September 2020).

21. Cheshire, J. Next Slide Please: Data Visualisation Expert on What's Wrong with the UK Government's Coronavirus Charts. The Conversation. 2020. Available online: https://theconversation.com/next-slide-please-data-visualisation-expert-on-whats-wrong-with-the-uk-governments-coronavirus-charts-149329 (accessed on 30 November 2020).

22. David, R. What's Wrong with COVID-19 Data Visualizations, and How to Fix It. Towards Data Science. 2020. Available online: https://towardsdatascience.com/whats-wrong-with-covid-19-data-visualizations-and-how-to-fix-it-3cdc9adc774d (accessed on 8 December 2020).

23. Sakai, R. What I Learned From COVID-19 Data Visualization. Nightingale. 2020. Available online: https://medium.com/nightingale/what-i-learned-from-covid-19-data-visualization-5c684eaa4698 (accessed on 8 December 2020).

24. Deahl, E. Better the Data You Know: Developing Youth Data Literacy in Schools and Informal Learning Environments. Master's Thesis, Massachusetts Institute of Technology, Department of Comparative Media Studies, Cambridge, MA, USA, 2014. [CrossRef]

25. Gebre, E.H. Young adults' understanding and use of data: Insights for fostering secondary school students' data literacy. *Can. J. Sci. Math. Technol. Educ.* **2018**, *18*, 330–341. [CrossRef]

26. Fontichiaro, K.; Oehrli, J.A.; Lennex, A. (Eds.) *Creating Data Literate Students*; Michigan Publishing: Ann Arbor, MI, USA, 2017.

27. Stiller, C.; Allmers, T.; Stockey, A.; Wilde, M. Statistical Literacy & Data Literacy—Grundbildung im Umgang mit empirischen Daten [Statistical Literacy & Data Literacy—Basic Education in empirical data use]. *Z. Schul-Und Prof.* **2020**, *2*, 144–160. [CrossRef]

28. Canipe, S. *The Basics of Data Literacy: Helping Your Students (and You!) Make Sense of Data*; NSTA Press: Arlington, VA, USA, 2015; pp. 70–71.

29. Mandinach, E.B.; Gummer, E.S. *Data Literacy for Educators Making It Count in Teacher Preparation and Practice*; Technology, Education-Connections the TEC Series; Teachers College Press: New York, NY, USA, 2016.

30. Mandinach, E.B.; Gummer, E.S. What does it mean for teachers to be data literate: Laying out the skills, knowledge, and dispositions. *Teach. Teach. Educ.* **2016**, *60*, 366–376. [CrossRef]

31. Mandinach, E.B.; Gummer, E.S. A systemic view of implementing data literacy in educator preparation. *Educ. Res.* **2013**, *42*, 30–37. [CrossRef]

32. Bowers, A.J. Quantitative Research Methods Training in Education Leadership and Administration Preparation Programs as Disciplined Inquiry for Building School Improvement Capacity. *J. Res. Leadersh. Educ.* **2017**, *12*, 72–96. [CrossRef]

33. Schüller, K.; Busch, P.; Hindinger, C. Future Skills: Ein Framework für Data Literacy—Kompetenzrahmen und Forschungsbericht. *Zenodo* **2019**. [CrossRef]

34. Heidrich, J.; Bauer, P.; Krupka, D. Future Skills: Ansätze Zur Vermittlung von Data Literacy in der Hochschulbildung [Future Skills: Approaches for Teaching Data Literacy in Higher Education]. *Zenodo* **2018**. [CrossRef]

35. Ridsdale, C.; Rothwell, J.; Smit, M.; Ali-Hassan, H.; Bliemel, M.; Irvine, D.; Kelley, D.; Matwin, S.; Wuetherick, B. *Strategies and Best Practices for Data Literacy Education: Knowledge Synthesis Report*; Technical Report; Dalhousie University: Halifax Regional Municipality, NS, Canada 2015. [CrossRef]

36. Abedjan, Z.; Brefeld, U.; Bürkle, J.; Desel, J.; Edlich, S.; Eppler, T.; Goedicke, M.; Heidrich, J.; Höppner, S.; Kast, S.; et al. *Data Science: Lern- und Ausbildungsinhalte [Data Science: Teaching and Educational Contents]*; GI—Gesellschaft für Informatik [German Informatics Society]: Bonn, Germany, 2019. Available online: https://gi.de/fileadmin/GI/Allgemein/PDF/GI_Arbeitspapier_Data-Science_2019-12_01.pdf (accessed on 6 October 2020).

37. Leonhardt, D. John Tukey, 85, Statistician; Coined the Word 'Software'. The New York Times 2000. Available online: https://www.nytimes.com/2000/07/28/us/john-tukey-85-statistician-coined-the-word-software.html (accessed on 17 September 2020).

38. Schumann, C.; Zschech, P.; Hilbert, A. Das aufstrebende Berufsbild des data scientist. *HMD Prax. Der Wirtsch.* **2016**, *53*, 453–466. [CrossRef]

39. Bhargava, R.; Deahl, E.; Letouzé, E.; Noonan, A.; Sangokoya, D.; Shoup, N. *Beyond Data Literacy: Reinventing Community Engagement and Empowerment in the Age of Data*; Technical Report; Data-POP Alliance and Internews and MIT Center for Civic Media and Harvard Hunanitarian Initiative and MIT Media Lab and ODI: New York, NY, USA, 2015.

40. Prado, J.C.; Ángel Marzal, M. Incorporating Data Literacy into Information Literacy Programs: Core Competencies and Contents. *Libri* **2013**, *63*, 123–134. [CrossRef]

41. Maltese, A.V.; Harsh, J.A.; Svetina, D. Data visualization literacy: Investigating data interpretation along the novice—Expert continuum. *J. Coll. Sci. Teach.* **2015**, *45*, 84–90. [CrossRef]

42. Beck, J.S.; Nunnaley, D. A continuum of data literacy for teaching. *Stud. Educ. Eval.* **2020**, 100871. [CrossRef]

43. Carlson, J.; Johnston, L.R. (Eds.) *Data Information Literacy: Librarians, Data, and the Education of a New Generation of Researchers*; Purdue Information Literacy Handbooks, Purdue University Press: West Lafayette, IN, USA, 2015. [CrossRef]

44. Smalheiser, N.R. *Data Literacy—How to Make Your Experiments Robust and Reproducible*; Elsevier Academic Press: Amsterdam, The Netherlands, 2017.

45. Herzog, D. *Data Literacy a User's Guide*; Sage: New York, NY, USA, 2016.

46. Weigend, A. *Data for the People How to Make Our Post-Privacy Economy Work for You*; Basic Books: New York, NY, USA, 2017.

47. Love, N.; Stiles, K.; Mundry, S.; DiRanna, K. *The Data Coach's Guide to Improving Learning for All Students Unleashing the Power of Collaborative Inquiry*; Corwin Press: Thousand Oaks, CA, USA, 2007.

48. Fontichiaro, K.; Lennex, A.; Hoff, T.; Hovinga, K.; Oehrli, J.A. (Eds.) *Data Literacy in the Real World: Conversations & Case Studies*; Michigan Publishing: Ann Arbor, MI, USA, 2017. [CrossRef]

49. Gemignani, Z.; Gemignani, C.; Galentino, R.; Schuermann, P. *Data Fluency: Empowering Your Organization with Effective Data Communication*; John Wiley & Sons, Inc.: Hoboken, NJ, USA, 2014.

50. Koltay, T. Data governance, data literacy and the management of data quality. *Int. Fed. Libr. Assoc. Ins. J.* **2016**, *42*, 303–312. [CrossRef]

51. Kuhn, S.; Kadioglu, D.; Deutsch, K.; Michl, S. Data Literacy in der Medizin: Welche Kompetenzen braucht ein Arzt? [Data literacy in medicine: Which competencies does a physician need?]. *Der Onkol.* **2018**, *24*, 368–377. [CrossRef]

52. Higgins, V.; Casasbuenas, V.; Ricard, J.; Carter, J. EmpoderaData: Data Litercy Assessment and Sustainable Development Goals Data Gaps. Technical Report, Data POP-Alliance and University of Manchester, 2019. Available online: https://datapopalliance.org/publications/empoderadata-data-literacy-assessment-and-sustainable-development-goals-data-gaps/ (accessed on 7 December 2020).

53. Schüller, K.; Busch, P. Data Literacy: Ein Systematic Review zu Begriffsdefinition, Kompetenzrahmen und Testinstrumenten. [Data Literacy: A Systematic Review of Definitions, Competency Frameworks and Test Instruments.] Available online: https://hochschulforumdigitalisierung.de/sites/default/files/dateien/HFD_AP_Nr_46_DALI_Systematic_Review_WEB.pdf (accessed on 21 January 2021).

54. Schuff, D. Data science for all: A university-wide course in data literacy. In *Analytics and Data Science: Advances in Research and Pedagogy*; Deokar, A.V., Gupta, A., Iyer, L.S., Jones, M.C., Eds.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 281–297._20. [CrossRef]

55. Cleveland, W.S. Data science: An action plan for expanding the technical areas of the field of statistics. *Int. Stat. Rev.* **2001**, *69*, 21–26. [CrossRef]

56. Kauermann, G. Data Science—Einige Gedanken aus Sicht eines Statistikers [Data science—Some thoughts by a statistician]. *Inform. Spektrum* **2020**, *42*, 387–393. [CrossRef]

57. Weihs, C.; Ickstadt, K., Ist Data Science mehr als Statistik? Ein Blick über den Tellerrand [Is data science more than statistics? A view beyond the horizon]. In *Faszination Statistik: Einblicke in Aktuelle Forschungsfragen und Erkenntnisse [Fascinating Statistics: A View on Current Research Questions and Insights]*; Krämer, W., Weihs, C., Eds.; Springer: Berlin/Heidelberg, Germany, 2019; pp. 203–210. [CrossRef]

58. Donoho, D. 50 years of data science. *J. Comput. Graph. Stat.* **2017**, *26*, 745–766. [CrossRef]

59. Bachelor Course Data Science—Recommended Order of Study Examination Regulations 2019. Available online: https://www.statistik.tu-dortmund.de/fileadmin/user_upload/Studium/Studiengaenge-Infos/Studienverlaufsplan_BSc_DaSc_2020.pdf (accessed on 17 September 2020).

60. Exemplary Study Plan Bachelor Data Science and Artificial Intelligence. Available online: https://www.uni-saarland.de/fileadmin/upload/studium/angebot/studplan/2019/Studplan_Data_vorl.pdf (accessed on 17 September 2020).

61. Data Science (BSc) Degree—Core Modules. Available online: https://warwick.ac.uk/study/undergraduate/courses-2020/datascience/ (accessed on 2 October 2020).

62. Bachelor Course Data Science—Requirements for Completing Lower and Upper Division of Data Science BA Major. Available online: https://data.berkeley.edu/academics/undergraduate-programs/data-science-programs/data-science-major (accessed on 27 November 2020).

63. Stockinger, K.; Stadelmann, T. Data Science für Lehre, Forschung und Praxis [Data science for teaching, research and practice]. *HMD Prax. Der Wirtsch.* **2014**, *51*, 469–479. [CrossRef]

64. Gausemeier, J.; Guggemos, M.; Kreimeyer, A. *Pilotphase Nationales Kompetenz-Monitoring (NKM): Bericht: Data Science—Auswahl, Beschreibung, Bewertung und Messung der Schlüsselkompetenzen für das Technologiefeld Data Science (Acatech DISKUSSION) [Pilot Stage*

National Competence Monitoring (NKM): Report: Data Science—Selection, Description, Evaluation and Measurement of Key Competencies for the Technological Field of Data Science (Acatech Discussion)]; Acatech: Munich, Germany, 2018.

65. Wrobel, S.; Voss, H.; Köhler, J.; Beyer, U.; Auer, S. Big data, big opportunities. *Inform. Spektrum* **2015**, *38*, 370–378. [CrossRef]
66. Kauermann, G. Data Science als Studiengang [Data Science as a Study Program]. In *Big Data: Chancen, Risiken, Entwicklungstendenzen [Big Data: Chances, Risks, Developments]*; König, C., Schröder, J., Wiegand, E., Eds.; Springer: Berlin/Heidelberg, Germany, 2018; pp. 87–95. [CrossRef]
67. Henderson, J.; Corry, M. Data literacy training and use for educational professionals. *J. Res. Innov. Teach. Learn.* **2020**. [CrossRef]
68. Sternkopf, H.; Mueller, R. Doing good with data: Development of a maturity model for data literacy in non-governmental organizations. In Proceedings of the 51st Hawaii International Conference on System Sciences, Waikoloa Village, HI, USA, 2–6 January 2018; pp. 5045–5054. [CrossRef]
69. Shields, M. Information literacy, statistical literacy, data literacy. *IASSIST Q.* **2005**, *28*, 6–11. [CrossRef]
70. Kalantzis, M.; Cope, B.; Chan, E.; Dalley-Trim, L. *Literacies*; Cambridge University Press: Cambridge, UK, 2016.
71. Mak, V.; Tjong, E.T.T.; Berlee, A. (Eds.) *Research Handbook in Data Science and Law*; Research Handbooks in Information Law, Edward Elgar Publishing: Cheltenham, UK, 2018.
72. Anderson, L.W.; Krathwohl, D.R.; Airasian, P.W.; Cruikshank, K.A.; Mayer, R.E.; Pintrich, P.R.; Raths, J.; Wittrock, M.C. *A Taxonomy for Learning, Teaching, and Assessing. A Revision of Bloom's Taxonomy of Educational Objectives*; Longman Publishing Group: London, UK, 2001.
73. Demchenko, Y.; Belloum, A.; Wiktorski, T. *EDISON Data science Framework: Part 1. Data Science Competence Framework (CF-DS)*, 2nd ed.; The EDISON Project: Fort Lauderdale, FL, USA, 2016. [CrossRef]
74. Demchenko, Y.; Manieri, A.; Wiktorski, T.; Bellowum, A. *EDISON Data Science Framework: Part 2. Data Science Body of Knowledge (DS-BoK)*, 2nd ed.; The EDISON Project: Fort Lauderdale, FL, USA, 2017. [CrossRef]
75. Lübcke, M.; Wannemacher, K. *Vermittlung von Datenkompetenzen an den Hochschulen: Studienangebote im Bereich Data Science [Teaching Data Competencies in Higher Education: Study Programs in Data Science]*; HIS-Institut für Hochschulentwicklung [HIS-Institute for University Development]: Hannover, Germany 2018.