Emerald Insight

Journal of Documentation
Data literacy: in search of a name and identity
Tibor Koltay

Article information:

Users who downloaded this article also downloaded:

For Authors

If you would like to write for this, or any other Emerald publication, then please use our Emerald for Authors service information about how to choose which publication to write for and submission guidelines are available for all. Please visit www.emeraldinsight.com/authors for more information.

About Emerald www.emeraldinsight.com

Emerald is a global publisher linking research and practice to the benefit of society. The company manages a portfolio of more than 290 journals and over 2,350 books and book series volumes, as well as providing an extensive range of online products and additional customer resources and services.

Emerald is both COUNTER 4 and TRANSFER compliant. The organization is a partner of the Committee on Publication Ethics (COPE) and also works with Portico and the LOCKSS initiative for digital archive preservation.

Data literacy: in search of a name and identity

Tibor Koltay

Department of Information and Library Studies, Szent István University, Jászberény, Hungary

Abstract

Purpose – The role of data literacy is discussed in the light of such activities as data a quality, data management, data curation, and data citation. The differing terms and their relationship to the most important literacies are examined. The paper aims to discuss these issues.

Design/methodology/approach – By stressing the importance of data literacy in fulfilling the mission of the contemporary academic library, the paper centres on information literacy, while the characteristics of other relevant literacies are also examined. The content of data literacy education is explained in the context of data-related activities.

Findings – It can be concluded that there is a need for data literacy and it is advantageous to have a unified terminology. Data literacy can be offered both to researchers, who need to become data literate science workers and have the goal to educate data management professionals. Several lists of competencies contain important skills and abilities, many of them indicating the close relationship between data literacy and information literacy. It is vital to take a critical stance on hopes and fears, related to the promises of widespread ability of (big) data.

Originality/value – The paper intends to be an add-on to the body of knowledge about information literacy and other literacies in the light of research data and data literacy.

Keywords Research, Data, Information literacy, Academic libraries

Paper type Viewpoint

Introduction

The increased interest in data characterizes not only the highly instrumented scientific and engineering research, but the social sciences and the humanities, as well. The vast amounts of data allow researchers to ask new questions in new ways, and – in the same time – also pose a wide range of concerns for access, management, and preservation (Borgman *et al.*, 2011). Data sharing is also an issue. To make data accessible, we need to develop appropriate technical and organizational infrastructures for storage and retrieval. Incentives and policies for researchers to share data are also indispensable (Kowalczyk and Shankar, 2011) and data literacy, which carries the potential of motivation, is one of the essential elements of this infrastructure.

Al in all, digital research data is a hot topic today. The professional literature of the 2010s clearly shows a tide in appearance of articles on digital research data issues.

For instance, the ACRL Research Planning and Review Committee (ACRL, 2012; ACRL, 2014) identified data curation (in a wide sense) as one of the top ten trends in academic libraries for both 2012 and 2014.

This paper describes the importance of data-related issues and activities in detail. We are also going to discuss, what role research data (including big data) plays today, then data literacy will be defined and its relationship to information literacy, academic literacy and other literacies will be explained. The content of data literacy education will also be analyzed.

Taking the recent tidal wave of interest in data-related issues, when discussing data literacy, we will rely mainly on literature from the 2010s.

Before trying to solve this naming problem and finding the proper identity of data literacy, we have to take a quick look at the role that the library can and should play in coping with the phenomenon of the "data deluge", as it is often called (Borgman, 2012; Little, 2012).

Research data and the library

In the same way as the library has traditionally facilitated access to documents, now information professionals could facilitate access to data, even though data do not necessarily fit into the same, document formats that libraries used to offer (Stuart, 2011). It is clear namely now that technology-driven research, known as eResearch, generates large quantities of data. However, the lack of tools, infrastructure, standardized processes and properly skilled personnel may impede the continued development of e-research (Carlson *et al.*, 2011).

No doubt, if libraries want to ensure that they remain relevant, services related to research data offer an opportunity to put the expertise of information professionals to good use (Christensen-Dalsgaard *et al.*, 2012). Realizing this, the Association of Research Libraries declared in 2012 that academic libraries are in a favourable position to help researchers to meet the challenges of a data-driven research paradigm because libraries and information professionals:

- have experience and skills in fostering cross-departmental, cross-campus, etc. communication and collaboration, required for effective research data management;

- are familiar with the research data needs of researchers and have been among the supporters of innovative publishing models, including open access publishing; and

- are increasingly providing data consultation services and information professionals are already involved in acquiring necessary abilities to manage data (Hswe and Holt, 2012).

The fields of involvement of academic libraries in providing research data services address the full data lifecycle, including planning, curation, and metadata creation and conversion (Tenopir *et al.*, 2011) as well as playing the critical role as data quality hubs on campus, by providing data quality auditing and verification services for the research communities (Giarlo, 2013).

The Association of European Research Libraries published recommendations for libraries to get started with research data management support that involves – among others – the development of metadata and data standards, creating data librarian (data specialists) posts, providing services for storage, discovery and permanent access and promoting data citation (Christensen-Dalsgaard *et al.*, 2012). Mooney and Silver (2011, p. 480) state that librarians "will not only need to learn how to apply traditional techniques of reference, instruction, and collection management to the unique format of data, but also how to raise awareness of this new service to the wider academic community".

According to Calzada Prado and Marzal (2013), academic libraries can have different responses to eScience (eResearch) and the growing need to use research data. They can hire specialized staff (data librarians or data specialists) or further data management and analysis training for (generally reference) librarians. Intensifying the collection of data sources and providing access to them including the participation in the

development of institutional data repositories is another option. Last, but not least they can incorporate data literacy in their instructional programs and services. Data literacy also appears in the predictions about future directions for academic libraries in the Library 2.0 world (Merrill, 2011).

Information professionals have begun to understand these needs. For instance, an online survey of 140 libraries in Australia, New Zealand, Ireland, and the UK identified needs in guidance in the handling and management of unpublished research data, including data literacy education (Corrall *et al.*, 2013).

In search of a name and identity

We must search for a name because the literature shows that differing concepts and terms are used. Besides data literacy there is data information literacy (Carlson *et al.*, 2011), science data literacy (Qin and D'Ignazio, 2010), and research data literacy (Schneider, 2013).

In our opinion, terminology should be unified around the concept and term of data literacy. We can adopt the definition of Calzada Prado and Marzal (2013), according to whom data literacy enables individuals to access, interpret, critically assess, manage, handle and ethically use data. Managing, figuring in this definition comprises preservation and curation. We can supplement the above with the viewpoint of Mandinach and Gummer (2013, p. 30), who define data literacy as "the ability to understand and use data effectively to inform decisions". They add that it is a specific skill set and knowledge base that enables us to transform data into information and ultimately into actionable knowledge.

When speaking about the issue of finding the proper name, it seems to be helpful to look at the difficulties experienced in the naming practices of information literacy. As Hunt (2004) explains, there may be no agreement on the precise definition of information literacy. Nonetheless, most people use the term information literacy rather than library instruction or information fluency. However, if we do not use the same language, it will be difficult to convince our stakeholders about the importance of information literacy education. In the case of data literacy we may experience the same.

Obviously, speaking about data literacy is impossible without delineating the concept of literacy. Without aiming and exhaustive discussion (done to a substantial extent by Bawden, 2001), we can affirm that while literacy and its counterpart, illiteracy were once well understood and well defined (Buschman, 2009), the growing role of digital technologies changed their meaning. It meant originally being knowledgeable with the body of writing of aesthetic merit, while now it often signifies being able to cope with the socio-technological changes and challenges brought by the convergence between media, telecommunication, information and communication technologies (Livingstone *et al.*, 2008).

Finding the identity of data literacy is based to a certain extent on its *raison d'être*. The perhaps most important objective is to exercise healthy criticism. We witness namely a widespread belief that the existence and access to research data, in particular to big data offe*r*s a higher form of intelligence and knowledge. There is an aura of truth, objectivity, and accuracy around it, as well. While big data is seen as a solution to many burning questions, it is regarded by many as a tool that threatens privacy, decreases civil freedoms, ushering increased state, and corporate control. The shifts to be expected of big data are probably more subtle than these, even though we cannot see this clearly among our current hopes and fears (Boyd and Crawford, 2012). This leads to the conclusion that one of the most important goals of data literacy education should be to foster critical thinking that keeps us away form the pitfalls of being overly

optimistic or unduly pessimistic, behaving in an excessively critical or uncritical way. In any case, critical assessment of data includes giving emphasis to the version of the given data set, the person responsible for it (ACRL, 2013). It also may involve understanding what data mean, including how to read graphs and charts appropriately, draw correct conclusions from data, and recognize when data is being used in misleading or inappropriate ways (Carlson *et al.*, 2011).

The content of data literacy education

No doubt, data literacy's learning outcomes need to be codified and (when acquired) best practices for data literacy programs have to be articulated (Hunt, 2004). Ascertaining these outcomes is the continuation of our efforts to delineate data literacy beyond a short definition. Several authors list the needed competencies, as a rule, using the words skills and abilities interchangeably and synonymously.

It seems to be unquestionable that an important characteristic of data literacy is the emphasis on the creation of data. Even though both Carlson *et al.* (2011) and Schneider (2013) use different terms (data information literacy and research data literacy, respectively) instead of data literacy, they stress the necessity to accommodate not only the data consumer's viewpoint, but the data producer's, as well.

It would be rather self-explanatory to take granted that different ways of using data is an important constituent of data literacy. This point of view gets emphasis, when we see that an examination of data literacy among teachers by Jacobs *et al.* (2009) shows that they define data as information that allows them to form opinions and make judgments, adding that data use requires sophisticated professional knowledge and ongoing attention to multiple sources of data (Jacobs *et al.*, 2009). This latter characteristic is underlined in regard to the preparation of educators (Mandinach and Gummer, 2013) and the most suitable one to be generalized for a use in other fields.

Data literacy education is not only a relatively new field for libraries, but it is one that cuts across disciplinary boundaries and across the traditional structures of academic library organizations. One sign of this is that humanities and social science scholars are in the process of becoming a new constituency for data literacy education (ACRL, 2013).

Data literacy education can have a dual purpose. One is rather self-explanatory, i.e. to achieve that students and researchers become data literate science workers. The other goal is to educate data management professionals (Qin and D'Ignazio, 2010; Schneider, 2013). In regard to the former, Haendel *et al.* (2012) speak about creating a culture of semantic scientists, stating that education in data literacy and information literacy should accompany scientific training to establish a new cultural standard, especially because researchers often do not realize that their own scholarly communications constitute a primary source of data. Digital humanities researchers may be discerned as a special subset of this group, because most digital humanities projects heavily rely on the interpretation of data and the field is struggling to maintain a healthy balance between the traditions and the ideals of humanistic endeavour, inherited from of long-established print scholarship and the effects of a growing digital infrastructure (Dalbello, 2011; Warwick, 2004). The move from text analysis to data mining and machine learning (Kirschenbaum, 2007) shows this. The idea of distant reading also seems to be influential. It is defined by Moretti (2005) as using graphs, maps, and trees instead of reading concrete, individual works, applying deliberate reduction and abstraction and concentrating on fewer elements that will allow us find sharper sense of their overall interconnection.

It is often difficult to separate data-related skills needed to become a successful researcher and to work as a data specialist first of all because part of these skills

and abilities frequently overlap each other. A list of the fields, where data literacy competencies can be used, clearly shows this:

- discovery and acquisition of data;

- data management;

- data conversion and interoperability (dealing with the risks and potential loss or corruption of information caused by changing data formats);

- metadata;

- data curation and re-use;

- data preservation;

- data analysis;

- data visualization; and

- ethics, including citation of data (Carlson *et al.*, 2011; RIN, 2011).

Data literacy needs to take into consideration the framework of future work skills and abilities, outlined by Davies *et al.* (2011). One of these abilities is the one to translate vast amounts of data into abstract concepts, as well as to understand data-based reasoning. This may lead to a situation, when we will start seeing almost everything that we come into contact, through the lens of data, and will regard it as computational, programmable.

Calzada Prado and Marzal (2013) emphasize the importance of knowing how to select and synthesize data and combine them with other information sources and prior knowledge. They also enumerate the following abilities:

- to identify the context in which data is produced and reused (data lifecycle);

- to recognize source data value, types and formats;

- to determine when data is needed;

- to access data sources appropriate to the information needed;

- to critically assess data and their sources;

- to determine and use suitable research methods;

- to handle and analyze data;

- to present quantitative information (specific data, tables, graphs, in reports and similar);

- to apply results to learning, decision making or problem-solving; and

- to plan, organize and self-assess throughout the process.

It is not by accident that context is mentioned in the above list in the first place. Context is utterly important. Dissociation of data from its context and the loss of context make reuse difficult, or impossible (Schneider, 2013).

From this list, we can single out four abilities as they practically repeat well-known abilities of information literacy, showing the close relationship between data literacy and information literacy. These are determining when data is needed, accessing sources appropriate to the information need, critically assessing data and their sources and applying results to learning, decision making or problem solving.

Accentuating data-driven decision making in education, Mandinach and Gummer (2013) set up a list of skills. They name knowing how to identify, collect, organize, analyze, summarize, and prioritize data. The last two skills are especially worth of attention as they the ones that do not appear in other lists. According to Mandinach and Gummer (2013), developing hypotheses, identifying problems, interpreting the data, and determining, planning, implementing, as well as monitoring courses of action also pertain to required skills and add the need for tailoring data literacy to the specific uses.

From the content of a series of instructional session on of socioeconomic data, described by Wong (2010), we can single out three aspects. One of them is exploring data evaluation and use. The other one is guiding students in understanding data-collection methods and dissemination channels. Introducing students to the rich varieties of data fulfilling different information needs should also be emphasized.

Data literate graduate students and researchers have to be familiar with some of the questions that data curation poses:

- Who owns the data?

- What requirements are imposed by others (e.g. funding agencies or publishers)?

- Which data should be retained?

- For how long should data be maintained?

- How should digital data be preserved?

- Are there ethical considerations?

- What sort of risk management is needed for research data?

- How are data accessed?

- How open should the data be?

- What alternatives to local data management exist? (Erway, 2013).

The need for reviewing data periodically is also worth of attention as it forms the basis of responsible decisions about disposal of "unnecessary" data, and preventing hardware or software obsolescence (Pryor, 2012).

A data literate researcher does not need to be a data curator. However, being familiar with the competencies of the curators and the fields, where the competencies, listed below, are of good use helps in developing data literacy:

- the ability to collaborate and work in teams;

- familiarity with scientific data sources;

- familiarity with quantitative research methods; and

- knowledge of general metadata standards (Si *et al.*, 2013).

From the fields, where these and other competencies can be used the following ones can be singled out:

- the data structure of different digital objects;

- the ways to assess the digital objects" authenticity, integrity and accuracy over time;

- storage and preservation policies, procedures and practices; and

- the risks of information loss or corruption of digital entities (Madrid, 2013).

Data-literate persons must be also aware of the limitations of computational thinking, mentioned earlier. This requires that they understand the limitations of data and remain able to act in the absence of data (Davies *et al.*, 2011).

Data literacy should include answers to the question about openness, asked above. It has to include the notion of open data, which has been advocated by a number of researchers in order to make science more open and accountable (Stuart, 2011). While doing this, we have to be aware of the legitimate boundaries of openness set by commercial interest, the protection of privacy, safety and security (The Royal Society, 2012).

Data quality is a driving force of data literacy (Carlson *et al.*, 2011; Madrid, 2013). It is determined by multiple factors. The first one is trust that depends on subjective judgments on authenticity, acceptability, or applicability of the data. Trust is also influenced by the given subject discipline, the reputation of those responsible for the creation of the data, and the biases of the persons who are evaluating it. The next factor is authenticity, which measures the extent to which the data is judged to represent the proper ways of conducting scientific research, including the reliability of the instruments used to gather the data, the soundness of underlying theoretical frameworks, the completeness, accuracy, and validity of the data. In order to evaluate authenticity, the data must be understandable. The condition for the evaluation of understandability of data is the presence of sufficient context in the form of documentation and metadata, and it requires the data be usable (Giarlo, 2013). If we want data to be usable, it has to be discoverable and accessible; and be in a usable file format. The individuals judging data quality need to have at their disposal an appropriate tool to access the data; which has to show sufficient integrity to be rendered. Integrity of data assumes that the data can be proven to be identical, at the bit level, to some previously accepted or verified state. Data integrity is required for usability, understandability and authenticity, thus it influences overall quality (Giarlo, 2013). Trust is related to cognitive authority, which has two levels. At an operational level, cognitive authority is the extent to which users think that they can trust the information. On a more general level, cognitive authority refers to "influences that a user would recognize as proper because the information therein is thought to be credible and worthy of belief" (Rieh, 2002, p. 146).

We already mentioned an essential aspect of data literacy. It is the profound understanding of the big data phenomenon, especially acknowledging that the decisive factor is not the quantity of data, but the capacity to search, aggregate, and cross-reference large data sets by virtue of the processing power of computers and networks. Data literate researchers and information professionals also have to know that data is no longer an exclusive issue for the sciences, but it is present in the social sciences, the humanities, arts and culture, as well (Boyd and Crawford, 2012).

It is vital not to forget about data citation, which allows the identification, retrieval, replication, and verification of data underlying published studies. Standardized forms of data citation could provide a motivation for researchers to share and publish their data, thus it has the potential to become a tool of reward and acknowledgment for them (Mooney and Newton, 2012). Unfortunately, despite numerous initiatives, at the moment, there is no standardization and consistency in data citation.

As service providers, data librarians should be acquainted with quantitative research methods, which enable them to process and analyze research data. To be able to provide support for researchers, they have to possess an extensive understanding of scientific data sources, which will enable them to recommending comprehensive and reliable data sources (Si *et al.*, 2013).

The summary of employers' requirements set against scientific data specialists (data librarians) by Si *et al.* (2013) shows high frequency of offering reference services for scientific research and data curation is regarded as one of the core duties of scientific data specialists, stating that in order to foster the data literacy of users, libraries require the instruction and training to users, including helping them understand the significance of scientific data curation and master the usage of various tools for data processing, data analysis and data statistics.

Data literacy and other literacies
The work of today's researcher mobilizes competencies on three levels:

- conceptual competencies that include among others innovative thinking, problem solving and critical thinking;

- human competencies, like social networking skills, self-management and cross-cultural interaction skills; and

- practical competencies that include media literacy and information literacy (Lee, 2013).

The perhaps best known form of practical competencies is information literacy. Education to information literacy emphasizes critical thinking and the necessity to recognize message quality. It has strong positions among literacies despite some (well founded) scepticism, highlighting the fact that this concept and especially the lack of information literacy has always seemed to be of more importance to academic librarians than to any other players in the information and education arena (Bawden and Robinson, 2009).

A broader interpretation of information literacy recognizes that the concept of information includes research data (RIN, 2011). Many writings, related to data literacy provide evidence of this. We already mentioned some. Besides of these, Fosmire and Miller (2008) speak about information literacy in the data world. Even though without referring to data literacy, Wang (2013) mentions that reference librarians frequently conduct information literacy sessions that educate the users about the existing data resources for their specific study areas. Calzada Prado and Marzal (2013) state that information literacy and data literacy form part of a scientific-investigative educational continuum, a gradual process of education that begins in school, is perfected and becomes specialized in higher education and becomes part of lifelong learning. In this later quality it is close to scientific literacy, which we will address shortly later.

Andretta *et al.* (2008) identified presenting, evaluating and interpreting qualitative and quantitative data as a learning outcome of information literacy. According to Hunt (2004), data literacy education should borrow heavily from information literacy education, even though the data literacy field is more fragmented than the domain of information literacy. Schneider (2013) also defines data literacy as a component of information literacy.

Si *et al.*, (2013) state that data-related services should be supported by professionals with both excellent scientific literacy and information literacy skills.

Adapting the ACRL information literacy competency standards for higher education (ACRL, 2000) to data information literacy, as done by Carlson *et al.* (2011), offers an interesting perspective. Obviously, we know that the ACRL standards have been criticized for representing a narrow and marginalized view of information literacy (Špiranec and Banek Zorica, 2010; Nazari and Webber, 2012). On the other hand the

standards not only acknowledge the value of research data, but include a number of explicit references to them (Calzada Prado and Marzal, 2013, Magnuson, 2013). Standard Three is especially relevant in this regard as it directs attention to evaluating information critically. If we translate this universal principle to data, we see that data information literate students are required not only to determine when and how to share data, but also document their own sources of data. This acquires significance in the light of the difficulty of documenting data, which is a fact that appeared in our writing several times. The reason for this is fairly simple. Critical examination of some information or data is practically impossible if we cannot rely on its documented feature, e.g. we cannot verify its provenance. Let us add that – as reflected by the definition, indicated above – it is an imperative for data literacy to be built on a critical approach towards data (Carlson *et al.*, 2011).

In a typical information literacy course we can assume the students know how to read, use a web browser and word processor. Data-related skills and abilities, even relatively simple ones, are not always so obviously present (Hunt, 2004). Despite of this, data literacy education could rely on the experiences of the writing across the curriculum movement, which presented a model since the 1970s for encouraging collaboration between faculty and librarians in developing composition and critical thinking skills (Bronshteyn and Baladad, 2006). It can provide information literacy with ways to resolve the increasingly irrelevant theory-practice divide as it places the student at the centre of the educational process (Elmborg, 2003).

Unfortunately, it would be utterly naïve presume that researchers easily (and readily) accept the need for acquiring data literacy skills. There is substantial evidence that the people in general as well as researchers hold themselves competent and skilful in dealing with information (Nicholas *et al.*, 2008). Data is probably not different in researchers' perception. Even more, they may feel more confident with it.

The relationship between data literacy and academic literacy has many faces. The latter has a grammatical dimension that information literate students must be taught and researchers must master. It involves the comprehension of the entire system of thinking, values, cultural identity and information flows of academia, which results in the ability to read, interpret, and produce texts valued in academia (Elmborg, 2006). Academic literacy is more closely associated with formal learning, than data literacy. This is especially true in higher education. Data literacy, affects the undergraduate student population to a lesser degree than academic literacy. Instead, data literacy has a true lifelong-learning character. The reason for this is in the fact that researchers in the past did not have to possess many of the data-related skills, at least not the ones related to data sharing and needed for citing data. In addition to this, the majority of the researchers have to acquire data literacy in the workplace, up to the point, when it becomes a standard literacy in graduate and doctoral education.

In his paper, Weideman (2003) enumerates some requirements set against academic literacy. Interpreting and using metaphors and idioms, or word play, rely mainly on linguistic abilities. The ability to understand academic vocabulary is different. It must be acquired together with the system of concepts and ways of thinking, characterizing a particular discipline or academia in general. Being aware of the logical development of an academic text is similarly two-faced. If we substitute data for information, we can identify a number of requirements that are relevant for data literacy. Sensitivity for meaning and for the intended audience, interpreting, using and producing information presented in graphic or visual format, as well as making distinctions between essential and non-essential information, fact and opinion, propositions and arguments; distinguishing

between cause and effect are examples of this. The requirement to classify, categorize and handle data and to do simple numerical estimations and computations are more directly connected to data literacy.

A number of other literacies can be mentioned and according to the type of connection to data literacy, there seem to be three slightly different types of these literacies.

The first type is close to data literacy by its topic. Statistical literacy pertains to this group. Schield (2004) is of the opinion that information literacy, statistical literacy and data literacy are tied together by a common set of problems and a similar level of approach. All three are more general than specific, they each involve interdisciplinary study and they deal with fundamentals. Besides of this, data literacy sets the overall context for evaluating the sources of data and is needed to access, manipulate and summarize the data, while statistical literacy guides that process.

Visual literacy is also often mentioned in relationship with data literacy because of the similarity between their topics. The ACRL Visual Literacy Competency Standards for Higher Education (ACRL, 2011) state the following:

Visual literacy is a set of abilities that enables an individual to effectively find, interpret, evaluate, use, and create images and visual media. Visual literacy skills equip a learner to understand and analyze the contextual, cultural, ethical, aesthetic, intellectual, and technical components involved in the production and use of visual materials. A visually literate individual is both a critical consumer of visual media and a competent contributor to a body of shared knowledge and culture.

Despite topical similarities, the most important points of this definition are the ones that show its relationship to information literacy and data literacy that can be called methodological. If we analyze it from this viewpoint, we can clearly see that critical approach remains prime interest. It is also reassuring and in accordance with modern approaches to information literacy that both consuming and producing receive emphasis. This is not different in relation to data literacy, either.

Data literacy is also connected to scientific literacy, which comprises methods, approaches, attitudes and skills, related to thinking scientifically and doing scientific research (National Academy of Sciences, 1996). This implies that the circle of potentially scientifically literate persons may be wide, even if only a small number of graduate students become scientists. Scientific literacy is not equivalent to teaching and learning sciences in public education. It is rather a complex set of knowledge of methods, approaches, attitudes and skills, related to a set of questions on how to do scientific research (Karvalics, 2013). The dominant similarity here is very much in the actors, i.e. in those persons, who need this literacy. Its connection to academic literacy, which was already mentioned as a main "neighbouring" literacy is almost self-explanatory.

Data literacy shares some distinguishing features with media literacy, especially in regard to the use and reuse of content in ways not imagined by the content creator (ACRL, 2013). Mentioning media literacy has two reasons. First, we know that there is a convergence among literacies caused by the (already mentioned) convergence between different forms of media and information and communication technologies (Livingstone et al., 2008). This provides a reason, why media literacy has relevance for a wide array of media, including research data. Second, we are aware of the fact that the capacity and interest in data-related issues is to a substantial extent result of the appearance of the Web 2.0. In this environment, users and their interests are represented in mediated spaces, which also serve as an environment to activate engagement with

others (Jarrett, 2008). This mediated nature influences data indirectly and to a lesser extent. Nonetheless, it requires that we approach data literacy not only as a manifestation of information literacy, but look at it through the lens of media literacy.

Media literacy comes into the picture as a relevant "general literacy". Obviously, information literacy can be labelled as a general one. The difference here is in the closeness of relationship between the two literacies, i.e. the influence of information literacy on data literacy is much more decisive than the one of media literacy.

Regarding "general literacies", digital literacy cannot be bypassed, either. As Qin and D'Ignazio (2010) put it, information literacy mobilizes the abilities and skills, related to finding, retrieving, analyzing, and using information. The same object, i.e. information, is aimed by digital literacy, which not only accentuates creating it (as we already mentioned this), but emphasizes the use of digital technology. When we underlined the importance of the data producer's viewpoint, we connected data literacy to the competences of digital literacy, which include "being comfortable with publishing and communicating information" (Bawden, 2008, p. 20).

Last, but not least, in virtue of its overarching nature, metaliteracy has to be mentioned. All the above literacies fit well into its framework, which provides the foundation for media literacy, digital literacy and other literacies, and emphasizes content (Mackey and Jacobson, 2011). With this the boundaries between information in information literacy from data in data literacy become blurred. In fact they never have been rigid, as information literacy has always been interested in the proper understanding and use of data that is converted into information (Schneider, 2013).

Conclusion

Data literacy is an evolving concept and it is in the need of using unified terminology. Data literacy education is an emerging field, which has a dual nature, i.e. its target audience is both potential data literate researchers and data specialists. The latter may either educate researchers to data literacy or fulfil different supporting functions in data-related research processes. Data literacy is akin to information literacy, while sharing a number of features with several other literacies. Among its features, borrowed from information literacy we find another duality. To be able to teach data literacy or be involved in any data-related research requires information professional to be data-literate themselves.

The success of data literacy will depend on how well we train information professional and make faculty and administrators understand why data literacy is imperative (Hunt, 2004). Re-skilling library staff is also utterly important in this process, as only a few libraries are able to hire new, specialized professionals (Christensen-Dalsgaard *et al.*, 2012).

Even though the spectrum of data-related activities is broad, information professionals need to be involved in them, centering on data citation and data literacy, emphasizing the educational nature and close relationship to information literacy and other relevant literacies of the latter.

References

ACRL (2000), *Information Literacy Competency Standards for Higher Education*, Association of College and Research Libraries, Chicago, IL.

ACRL (2012), "ACRL research planning and review committee. Top ten trends in academic libraries. a review of the trends and issues affecting academic libraries in higher education", *College & Research Libraries News*, Vol. 73 No. 6, pp. 311-320.

ACRL (2013), *Intersections of Scholarly Communication and Information Literacy: Creating Strategic Collaborations for a Changing Academic Environment*, Association of College and Research Libraries, Chicago, IL.

ACRL (2014), "ACRL research planning and review committee. Top ten trends in academic libraries. A review of the trends and issues affecting academic libraries in higher education", *College & Research Libraries News*, Vol. 75 No. 6, pp. 294-302.

Andretta, S., Pope, A. and Walton, G. (2008), "Information literacy education in the UK", *Communications in Information Literacy*, Vol. 2 No. 1, pp. 36-51.

Bawden, D. (2001), "Information and digital literacies: a review of concepts", *Journal of Documentation*, Vol. 57 No. 2, pp. 218-259.

Bawden, D. (2008), "Origins and concepts of digital literacy", in Lankshear, C. and Knobel, M. (Eds), *Digital Literacies: Concepts, Policies and Practices*, edited by New York, Peter Lang, NY, pp. 17-32.

Bawden, D. and Robinson, L. (2009), "The dark side of information: overload, anxiety and other paradoxes and pathologies", *Journal of Information Science*, Vol. 35 No. 2, pp. 180-191.

Borgman, C.L. (2012), "The conundrum of sharing research data", *Journal of the American Society for Information Science and Technology*, Vol. 63 No. 6, pp. 1059-1078.

Borgman, C.L. (Ed.) (2011), *Fostering Learning in the Networked World: The Cyberlearning Opportunity and Challenge*, DIANE Publishing, Philadelphia, PA.

Boyd, D. and Crawford, K. (2012), "Critical questions for big data: provocations for a cultural, technological, and scholarly phenomenon", *Information, Communication & Society*, Vol. 15 No. 5, pp. 662-679.

Bronshteyn, K. and Baladad, R. (2006), "Librarians as writing instructors: using paraphrasing exercises to teach beginning information literacy students", *Journal of Academic Librarianship*, Vol. 32 No. 5, pp. 533-536.

Buschman, J. (2009), "Information literacy, new literacies, and literacy", *Library Quarterly*, Vol. 79 No. 1, pp. 95-118.

Calzada Prado, J. and Marzal, M.Á. (2013), "Incorporating data literacy into information literacy programs: core competencies and contents", *Libri*, Vol. 63 No. 2, pp. 123-134.

Carlson, J., Fosmire, M., Miller, C.C. and Nelson, M.S. (2011), "Determining data information literacy needs: a study of students and research faculty", *Portal: Libraries and the Academy*, Vol. 11 No. 2, pp. 629-657.

Christensen-Dalsgaard, B., van den Berg, M., Grim, R., Horstmann, W., Jansen, D., Pollard, T. and Roos, A. (2012), "Ten recommendations for libraries to get started with research data management", LIBER, available at: www.libereurope.eu/news/ten-recommendations-for-libraries-to-get-started-with-research-data-management (accessed 18 June 2014).

Corrall, S., Kennan, M.A. and Afzal, W. (2013), "Bibliometrics and research data management services: emerging trends in library support for research", *Library Trends*, Vol. 61 No. 3, pp. 636-674.

Dalbello, M. (2011), "A genealogy of digital humanities", *Journal of Documentation*, Vol. 67 No. 3, pp. 480-506.

Davies, A., Fidler, D. and Gorbis, M. (2011), *Future Work Skills, 2020*, Institute for the Future, Palo Alto, CA, available at: www.iftf.org/our-work/global-landscape/work/future-work-skills-2020/ (accessed 18 June 2014).

Elmborg, J. (2006), "Critical information literacy: implications for instructional practice", *Journal of Academic Librarianship*, Vol. 32 No. 2, pp. 192-199.

Elmborg, J.K. (2003), "Information literacy and writing across the curriculum: sharing the vision", *Reference Services Review*, Vol. 31 No. 1, pp. 68-80.

Erway, R. (2013), "Starting the conversation: university-wide research data management policy", *Educause Review Online*, December 6, available at: www.educause.edu/ero/article/starting-conversation-university-wide-research-data-management-policy (accessed 18 June 2014).

Fosmire, M. and Miller, C. (2008), "Creating a culture of data integration and interoperability: librarians and earth science faculty collaborate on a geoinformatics course", Paper No. 16, *Proceedings of the IATUL Conferences*, available at: http://docs.lib.purdue.edu/iatul/2008/papers/16 (accessed 18 June 2014).

Giarlo, M. (2013), "Academic libraries as quality hubs", *Journal of Librarianship and Scholarly Communication*, Vol. 1 No. 3, pp. 1-10.

Haendel, M.A., Vasilevsky, N.A. and Wirz, J.A. (2012), "Dealing with data: a case study on information and data management literacy", *PLoS Biology*, Vol. 10 No. 5, p. e1001339.

Hswe, P. and Holt, A. (2012), *A New Leadership Role for Libraries*, Association of Research Libraries, Washington, DC, available at: http://old.arl.org/rtl/eresearch/escien/nsf/leadershiproles.shtml (accessed 18 June 2014).

Hunt, K. (2004), "The challenges of integrating data literacy into the curriculum in an undergraduate institution", *IASSIST Quarterly*, Vol. 28 No. 2, pp. 12-15, available at: www.iassistdata.org/downloads/iqvol282_3hunt.pdf (accessed 18 June 2014).

Jacobs, J., Gregory, A., Hoppey, D. and Yendol-Hoppey, D. (2009), "Data literacy: understanding teachers' data use in a context of accountability and response to intervention", *Action in Teacher Education*, Vol. 31 No. 3, pp. 41-55.

Jarrett, K. (2008), "Interactivity is evil! a critical investigation of web 2.0", *First Monday*, Vol. 13 No. 3, p. 13, available at: http://firstmonday.org/htbin/cgiwrap/bin/ojs/index.php/fm/article/viewArticle/2140/1947 -(accessed 18 June 2014).

Karvalics, Z.L. (2013), "From scientific literacy to lifelong research: a social innovation approach", in Kurbanoğlu, S. *et al.* (Eds), *Worldwide Commonalities and Challenges in Information Literacy Research and Practice*, Springer International, Cham, pp. 126-133.

Kirschenbaum, M. (2007), "The remaking of reading: data mining and the digital humanities", *NGDM 07: National Science Foundation Symposium on Next Generation of Data Mining and Cyber-Enabled Discovery for Innovation*, Baltimore, MD, 10-12 October, available at: www.csee.umbc.edu/~hillol/NGDM07/abstracts/talks/MKirschenbaum.pdf (accessed 18 June 2014).

Kowalczyk, S. and Shankar, K. (2011), "Data sharing in the sciences", *Annual Review of Information Science and Technology*, Vol. 45 No. 1, pp. 247-294.

Lee, A.Y. (2013), "Literacy and competencies required to participate in knowledge societies", *Conceptual Relationship of Information Literacy and Media Literacy in Knowledge Societies*, UNESCO, Paris.

Little, G. (2012), "Managing the data deluge", *Journal of Academic Librarianship*, Vol. 38 No. 5, pp. 263-264.

Livingstone, S., van Couvering, E.J. and Thumin, N. (2008), "Converging traditions of research on media and information literacies: disciplinary and methodological issues", in Leu, D.J. *et al.* (Eds), *Handbook of Research on New Literacies*, Lawrence Erlbaum Associates, Hillsdale, NJ, pp. 103-132.

Mackey, T. and Jacobson, T. (2011), "Reframing information literacy as a metaliteracy", *College & Research Libraries*, Vol. 72 No. 1, pp. 62-78.

Madrid, M.M. (2013), "A study of digital curator competences: a survey of experts", *International Information & Library Review*, Vol. 45 Nos 3-4, pp. 149-156.

Magnuson, M.L. (2013), "Web 2.0 and information literacy instruction: aligning technology with ACRL standards", *Journal of Academic Librarianship*, Vol. 39 No. 3, pp. 244-251.

Mandinach, E.B. and Gummer, E.S. (2013), "A systemic view of implementing data literacy in educator preparation", *Educational Researcher*, Vol. 42 No 1, pp. 30-37.

Merrill, A. (2011), "Library+", *Public Services Quarterly*, Vol. 7 Nos 3/4, pp. 144-148.

Mooney, H. and Silver, B. (2010), "Spread the news: promoting data services", *College & Research Libraries News*, Vol. 71 No 9, pp. 480-483.

Mooney, H. and Newton, M.P. (2012), "The anatomy of a data citation: discovery, reuse, and credit", *Journal of Librarianship and Scholarly Communication*, Vol. 1 No 1, pp. 1-14.

Moretti, F. (2005), *Graphs, Maps, Trees: Abstract Models for Literary Theory*, Verso, London.

National Academy of Sciences (1996), *National Science Education Standards*, National Academy Press, Washington, DC.

Nazari, M. and Webber, S. (2012), "Loss of faith in the origins of information literacy in e-environments: proposal of a holistic approach", *Journal of Librarianship and Information Science*, Vol. 44 No. 2, pp. 97-107.

Nicholas, D., Huntington, P., Jamali, H.R. and Dobrowolski, T. (2008), "The information-seeking behaviour of the digital consumer: case study the virtual scholar", in Nicholas, D. and Rowlands, I. (Eds), *Digital Consumers: Reshaping the Information Professions*, Facet, London, pp. 113-58.

Pryor, G. (2012), "Why manage research data?", in Pryor, G. (Ed.), *Managing Research Data*, Facet, London, pp. 1-16.

Qin, J. and D'Ignazio, J. (2010), "Lessons learned from a two-year experience in science data literacy education", *Proceedings of the 31st Annual IATUL Conference, 20-24 June*, available at: http://docs.lib.purdue.edu/iatul2010/conf/day2/5 (accessed 18 June 2014).

Rieh, S.Y. (2002), "Judgment of information quality and cognitive authority in the web", *Journal of the American Society for Information Science and Technology*, Vol. 53 No 2, pp. 145-161.

RIN (2011), *The Role of Research Supervisors in Information Literacy*, Research Information Network, London, available at: www.rin.ac.uk/system/files/attachments/Research_supervisors_report_for_screen.pdf (accessed 18 June 2014).

Schield, M. (2004), "Information literacy, statistical literacy and data literacy", *IASSIST Quarterly*, Vol. 28 Nos 2/3, pp. 6-11.

Schneider, R. (2013), "Research data literacy", in Kurbanoglu, S. *et al.* (Ed.), *Worldwide Commonalities and Challenges in Informbration Literacy Research and Practice*, Springer International, Cham, pp. 134-140.

Si, L. *et al.* (2013), "The cultivation of scientific data specialists: development of LIS education oriented to e-science service requirements", *Library Hi Tech*, Vol. 31 No 4, pp. 700-724.

Špiranec, S. and Banek Zorica, M. (2010), "Information literacy 2.0: hype or discourse refinement?", *Journal of Documentation*, Vol. 66 No. 1, pp. 140-153.

Stuart, D. (2011), *Facilitating Access to the Web of Data*, Facet, London.

Tenopir C. *et al.* (2011), "Data sharing by scientists: practices and perceptions", *PLoS ONE*, Vol. 6 No. 6, available at: www.plosone.org/article/info:doi/10.1371/journal.pone.0021101 (accessed 18 June 2014).

The Royal Society (2012), *Science as an Open Enterprise*, The Royal Society, London, available at: https://royalsociety.org/uploadedFiles/Royal_Society_Content/policy/projects/sape/2012-06-20-SAOE.pdf (accessed 20 February 2014).

Wang, M. (2013), "Supporting the research process through expanded library data services", *Program*, Vol. 47 No. 3, pp. 282-303.

Warwick, C. (2004), "Print scholarship and digital resources", in Schreibman, S., Siemens, R. and Unsworth, J. (Eds), *A Companion to Digital Humanities*, Blackwell, Oxford, pp. 366-382.

Weideman, A. (2003), "Assessing and developing academic literacy", *Per linguam*, Vol. 19 Nos 1/2, pp. 55-65.

Wong, G.K.W. (2010), "Facilitating students' intellectual growth in information literacy teaching", *Reference and User Services Quarterly*, Vol. 50 No. 2, pp. 114-118.

415

Further reading

Beeson, I. (2005), "Judging relevance: a problem for e-literacy", *ITALICS*, Vol. 4 No. 2, pp. 210-219, available at: www.ics.heacademy.ac.uk/italics/vol5iss4/beeson.pdf (accessed 18 June 2014).

Federer, L. (2013), "The librarian as research informationist: a case study", *Journal of the Medical Library Association*, Vol. 101 No. 4, pp. 298-302.

Mooney, H. and Silver, B. (2010), "Spread the news: promoting data services", *College & Research Libraries News*, Vol. 71 No. 9, pp. 480-483.

Corresponding author
Professor Tibor Koltay can be contacted at: koltay.tibor@abpk.szie.hu

This article has been cited by:

1. C. BasiliInformation Literacy Requirements for Open Science 229-248. [CrossRef]
2. Tibor Koltay Szent István University, Jászberény, Hungary . 2016. Are you ready? Tasks and roles for academic libraries in supporting Research 2.0. *New Library World* **117**:1/2, 94-104. [Abstract] [Full Text] [PDF]
3. Tibor Koltay, Sonja ŠpiranecLibraries Meet Research 2.0: 32-52. [CrossRef]