June 2016

The Physical, Mental and Social Dimensions of Documents

Michael Buckland

School of Information, University of California, Berkeley, buckland@ischool.berkeley.edu

Introduction

In the development of the program in Documentation Studies at the University of Tromsø, Norway, in 1996, one of the guiding principles was that

> ". . . one should view the document from three complementary angles: physical, social, and mental, in combination enabling a complete description. This does not mean that the document possesses each of these three features to some degree but that it can be viewed simultaneously as a physical, social, and mental phenomenon. From this perspective, the core issue is how these dimensions interact with each other in different ways" (Lund 2009, 424).

These three angles are summarized by Olsen, Lund, Ellingsen, and Hartvigsen (2012, 111-113) and Skare (2009) has explained that the angles complement each other, but not in a mutually exclusive way. Here I discuss the "core issue" of how these three dimensions interact with each other.

The Physical

A document is some entity regarded by someone as signifying something. It has to be a physical, material entity unless and until we want to expand into extrasensory perception, direct divine inspiration, or telepathy. It is sometimes assumed or implied that electronic records ("the virtual") are somehow not physical, but this is an error because electronic systems are physical. They do not achieve much without, for example, magnetic charges or electrical power.

One can discuss a *text* or a *work* in an abstract sense but texts and works can exist as documents only in some physical manifestation. Information systems are supposed to inform people but they do so always and only through physical stuff. All engineered information systems operate on physical records whether print on paper, holes in a punch card, magnetized bits, optical pulses, or other physical media.

The physical aspect means that all documents exist in space and time. The spatial aspect means that all documents occupy physical space somewhere and anything existing in physical space can, in principle, be moved to a new location, though ease of mobility varies greatly. The temporal aspect of documents is also significant. It may take time to read a text or hear a recording. Some kinds of documents are designed to change over time, for example moving images media and performances.

And, as time passes, anything physical will change sooner or later, making stability and preservation important practical issues. An extreme case is the vulnerability of electronic records to loss or corruption. Religious rituals and opera performance, for example, both potentially very meaningful, ordinarily involve both movement in space and the passing of time.

The history of document technology – writing, printing, telecommunications, copying -- can be seen as a continuing effort to reduce the constraints of time and place (Buckland 2015).

The Mental

The physical dimension is a necessary but not a sufficient condition for being a document. Someone must view it as signifying (or potentially signifying) something, even if unsure of what the significance might be. Suzanne Briet, in her explanation of what is a document, stated that it would have to be considered *evidence*: "A document is proof in support of a fact" (Briet, 1951/2006, 9). Her French original used the French word *preuve*, which corresponds to the English *proof*, but can also refer to *testimony* and *evidence*).

Status as a document (as actually or potentially evidence of something) is an individual, personal mental judgment and, therefore, subjective. Such a perception occurs only in a living mind and, with any living, learning mind, the perception can change as what the individual knows changes, as it does continually until death. Although the consequences of this perception might be observable, the perception itself is neither observable nor measurable.

The Social

The adjective "social" is widely used in relation to documents. We read about "the social life of documents" (e.g. Brown & Duguid, 2000) or of documents as "social traces" (e.g. Ferraris, 2013). But if we assume that only an individual can be informed by a document (through a mental construction) then caution is needed to distinguish the *social* from the *mental*. If we set aside the use of *social* when used figuratively to denote a multiplicity of individuals engaged in subjective mental activity as belonging more properly to the *mental* aspect, the *social* can include the sociology of knowledge, especially interactions between two or more different individuals influencing each other in their understanding of reality. (For a convenient introduction see Zerubavel (1997); also Mannheim (1936, chap. 1) and Berger & Luckmann, 1966).

A central concept in the sociology of knowledge is intersubjectivity. An individual can make a subjective idea objectively perceptible by others. For example, a hostile attitude may be made objective by a frown, by the threatening use of a weapon, or by using words, to another individual who then makes a subjective interpretation and, probably, react accordingly. In this way, subjective understandings develop among two or more individuals in a related, dialectic way. These more or less shared subjective understandings – *intersubjective* understandings – form the basis of the shared culture of any social group. The multiplicity, complexity, and fluidity of social groupings needs to be noted.

The social dimension is reflected in collaborative actions, such as teamwork and joint coercion.

Combinations

A few examples can illustrate combinations of these three dimensions.

The physical and the social dimensions

A text may be authored through the mental efforts of a solitary individual but physical documents are ordinarily the result of the actions of many different people. A printed book depends on paper manufacturers, printers, publishers, typesetters, binders, book

retailers, and many others. Shared financial, transportation, and other infrastructures support all of their varied contributions, and a book would not be printed in the absence of readers. (See McGann, 1983; and McKenzie, 1986).

All communities depend on the division of social labor resulting in a social division of specialized knowledge and, increasingly, members' dependence on second-hand knowledge (Wilson, 1993). However, the division of labor can only operate with coordination, whether political, through management, or economic, through markets. Coordination requires communication which in practice means documents. So the rise of the so-called "information society" could be more accurately be described as the rise of a "document society".

It is the rise of physical documentary techniques such as writing, printing, telecommunications, copying, and computing that has enabled the social division of labor and what is ordinarily meant by "information society".

The social and the physical combine in ways that involve the mental dimension less directly in the area of information policy in which social powers are used to enable or, commonly, to restrict mental activity through economic, legislative, political, and other means. Examples include the regulations governing intellectual property, textbook adoption, privacy, libel, technical standards, and national security. These affordances influence mental activity indirectly by influencing the opportunities.

The social and the mental dimensions

Behavior derives from both nature and nurture. Our mental behavior is profoundly influenced by nurture, by what we learn directly or indirectly from others. Nurture is a social process. Our culture and cultural heritage are socially derived. As Ludwik Fleck (1935/1979; see also Cohen & Schnelle, 1986) emphasized, understanding a written text requires taking into account the writer's cultural context. In terms of our present discussion, a document must have both physical and mental properties, but since the mental processes are culturally entangled with the social, the status of being a document necessarily also entails a social dimension indirectly through the mental. This alone is sufficient justification for insisting that every document must necessarily have a social angle as well as mental and physical angles.

Discussion

We have so far focused on pairs of dimensions, but it can be noticed that the third dimension sooner or later emerges as implicated. We use, and need to use, documents to aid, to persuade, to control, and in many other ways and in doing so the three angles – the social, the physical and the mental – are all directly in use. This can be illustrated with three examples.

Infrastructure

The production, dissemination, and accessibility of documents are enabled (and constrained) by infrastructure that is socially provided, including legal regimes underlying commerce and intellectual property, standardized terminology in metadata, markets, subsidies, and restrictions relating to decency, privacy, security and other

cultural values. In brief, the opportunities for mental engagement with (physical) documents is heavily framed by social forces.

Relevance

Information services are purposive, expected to serve someone's mental activity needs beneficially and a document is said to be relevant if useful. Since the 1960s *relevance* has been regarded as central to information science and made the primary basis for the quantitative evaluation of information retrieval systems. But, despite sustained attention by many talented minds, relevance has resisted satisfying definition or measurement. Howard White (2010) provides an excellent account of relevance theory. He states, correctly, that although relevance is well understood, it resists satisfying definition, observation, or scientific treatment, as was noted by early critics (e.g. Mortimer Taube's denunciation of "the pseudo-mathematics of relevance" (Taube, 1965)).

To be relevant a document must be useful to an actual human being's mental activity and is, therefore, idiosyncratic, hard to predict, and unstable. (Relevance to a specific need of a specific person is sometimes named pertinence.) Ordinarily one can only make a judicious guess that a given document is likely to be relevant to a given query for a supposed population of users at some point in time.

There are many difficulties associated with relevance which we need not consider here. An explanation of the basic problem is that documents have both physical and mental aspects. Scientific measurement depends on there being something physical to measure. The physical aspects of documents can be measured and so treated scientifically, but the highly situational, unstable, idiosyncratic, and subjective mental angle cannot. Because every document also has a significant but inaccessible mental aspect its relevance cannot be measured scientifically. For this reason relevance can never be satisfactorily a scientific matter in the normative sense of formal and physical sciences such as mathematics and physics. In practice we fall back on distant substitutes. We can use the physical angle only, primarily of coded character strings and use character strings in a query to discover similar character strings in documents that might be discourse on the same topic. (We should not say words, because character strings do not distinguish different words spelled the same.) The matching of character strings works quite well but not very reliably. We can ask a jury to predict whether a document is likely to be relevant to a hypothetical inquirer and we can ask an inquirer, after a search, whether a document was relevant, but either judgement might not be valid for someone else or for the same person at another time.

A scientific approach to relevance could work very well if a document had only a physical aspect and not also a mental one. We see this situation in the case of the modelling of signaling reliability developed by Claude Shannon as Communication Theory and now better known as Information Theory. The scientific quality and practical utility of this model is beyond question and it can be achieved because no mental or social properties are involved, only physical properties. (I thank Wayne de Fremery for this insight.) A desire to make this Information Theory a central component in Library and Information Science has not proven successful and the reason is not hard to see. For any Information Science concerned with what individuals know requires a mental angle and Shannon-Weaver Information Theory is powerful precisely because the mental angle

is absent. It can be useful as a tool, just as queuing theory and other quantitative tools can be, but despite its name it cannot claim any greater special status.

Language and documents

Fifty years ago Berger and Luckmann (1966) in their *The social construction of reality: A treatise on the sociology of knowledge* provided a detailed explanation of how the subjective can be made objective, and thereby accessible to others, through an expression (a frown), a gesture (with a dagger), or a conversation. They rightly emphasize the power of language, but in doing so an opportunity was lost in what could have been added. Language's importance is as an ingredient in communication and is largely, and increasingly, expressed in documents. Had that point been made the study of documents and of documentation, it might have received much more attention in the past half-century.

Conclusion

Two claims have been examined: the claim that any and every document has a physical angle *and* a mental angle *and* a social angle and the related claim that in considering documents none of these three angles can be completely understood without acknowledging the other two. Both claims find justification.

Any document must necessarily be physical, but physicality alone is not a sufficient condition. There must also be a mental angle for a physical entity to be a considered a document. Since the mental angle is different in kind from the physical and since only the physical aspect can be adequately treated in traditionally scientific methods and so scientific solutions are necessarily incomplete. This explains the inherent difficulties in the use of *relevance* in the evaluation of retrieval systems and illuminates the contrast between *relevance*, which has a mental angle, and Shannon-Weaver Information theory which does not.

The social angle is necessarily implicated in document theory because mental activity is influenced by cultural nurture and also, in practice, because the disposition of (physical) documents is influenced by social controls.

It is clear that there is plenty of scope for examination of the "core issue" of documentation in the coming twenty years.

References

Berger, P. L. & T. Luckmann. 1966. The social construction of reality: a treatise in the sociology of knowledge. Garden City, NY: Doubleday.

Briet, S. (1951/2006). What is documentation? English translation of the classic French text. Lanham, MD: Scarecrow Press. Also http://info.slis.indiana.edu/~roday/what%20is%20documentation.pdf

Brown, J. S. & P. Duguid. 2000. The social life of information. Boston: Harvard Business School Pr.

Buckland, M. K. (2015). Document theory: An introduction. In M. Willer, A. J. Gilliland & M. Tomić (Eds), Records, Archives and Memory: Selected Papers from the Conference and School on Records, Archives and Memory Studies, University of Zadar, Croatia, May 2013 (pp. 223-237). Zadar: University of Zadar, 2015. ISBN

978-953-331-080-0. Preprint at
http://people.ischool.berkeley.edu/~buckland/zadardoctheory.pdf

Cohen, R. S. & T. Schnelle (Eds). (1986). Cognition and fact: Materials on Ludwik Fleck. (Boston Studies in the Philosophy of Science 87). Dordrecht: Reidel Publishing Co.

Ferraris, M. (2013). Documentality: why it is necessary to leave traces. New York : Fordham University Pr.

Fleck, L. (1935/1979). Genesis and development of a scientific fact. Chicago: Univ. of Chicago Press.

Lund, N. W. (2009). Document theory. Annual Review of Information Science and Technology, 43, 399-432.

Mannheim, K. 1936. Ideology and utopia; an introduction to the sociology of knowledge. New York, Harcourt, Brace & World

McGann, J. (1983). A critique of modern textual criticism. Chicago : University of Chicago Pr,

McKenzie, D. F. (1986). Bibliography and the sociology of texts. London: British Library.

Olsen, B. I., N. W. Lund, G. Ellingsen, & G. Hartvigsen. (2012). Document theory for the design of socio-technical systems: A document model as ontology of human expression. Journal of Documentation, 68 (1), 100-126.

Skare, R. (2009) Complementarity – a concept for document analysis? Journal of Documentation, 65(5), 834 - 840

Taube, M. (1965). A note on the pseudo-mathematics of relevance. American Documentation, 16 (2), 69-72.

White, H. D. (2010). Relevance in theory. In: M. Bates (Ed.), Encyclopedia of library and information sciences, 3rd ed. (pp 4498-4511). Boca Raton, FL : CRC Press.

Wilson, P. (1983). Second-hand knowledge: an inquiry into cognitive authority. Westport, CT: Greenwood Pr.

Zerubavel, E. (1997). Social mindscapes: an invitation to cognitive sociology. Cambridge, MA: Harvard University Pr.