CITY UNIVERSITY LONDON

A Library Love Triangle? : An analysis of the relationship between data, information and knowledge in Library and Information Studies

(or, Pullman's Dust : A new model for data, information and knowledge).

James Atkinson

January 2016

Submitted in partial fulfilment of the requirements for the degree of MSc in Library Science.

Supervisor: Lyn Robinson

Abstract

Across the decades, many writers within the Library and Information Science (LIS) sector have sought to define and discuss issues surrounding data, information and knowledge, often drawing on other disciplines. What is lacking, however is any kind of consensus as different ideas and opinions are merely presented and argued.

This dissertation seeks to compare and contrast different definitions of the concepts data, information and knowledge within LIS before comparing and discussing different views of the relationships between them. It will also look for evidence within three fields of LIS other than Information Theory (Information Retrieval, Information Literacy and Knowledge Management) in order to see if, and how, the more practical side of the profession alters the view of the theoretical.

The Methodology is one of desk research, a literature search conducted in order to ensure a comprehensive and wide-ranging discussion.

The dissertation will find that there is little agreement as to how data, information and knowledge interact, and that this is due to different views of the definitions of the concepts, different views of the LIS discipline and because of the widely differing views of how knowledge interacts with information.

However, it will be found that knowledge is a large concept comprising of many parts and that can be affected in many ways; and that knowledge should be viewed as a more overarching concept than the many linear views of the relationship between data, information and knowledge often shown.

As such, it will be suggested that Pullman's Dust (a concept drawn from Philip Pullman's *His Dark Materials* books) provides an adequate analogy for knowledge and a new model of the relationship between data, information and knowledge is suggested along with other future suggestions for research.

Table of Contents

List of Figures

Acknowledgements

I would like to thank my supervisor, Lyn Robinson, for her help and support.

Also, my work colleagues who are former #citylis, Oskar Brown and Kathryn Drumm, especially, for their advice and support.

I am additionally grateful to the other members of City University London Library's Information Literacy Group. Their support in introducing me to the world of Information Literacy has been invaluable to this dissertation. In particular I would like to thank Claire Packham, Alex Asman, Catherine Radbourne and Fiona Paterson with whom I have undertaken project work.

Finally, I would like to thank my wife, Priya Bassan, for her love and support since I began my #citylis journey shortly after our honeymoon in 2013. In particular, for proofreading what follows and putting up with me studying through a third Christmas in a row.

"The nature of the relations among them is debateable, as well as their meanings." - Chaim Zins (2007a)

1. Introduction

Across the decades, many writers within the Library and Information Science (LIS) sector have sought to define and discuss issues surrounding data, information and knowledge, often drawing on other disciplines. What is lacking, however is any kind of consensus as different ideas and opinions are merely presented and argued.

Information theory has been studied since the Second World War in different ways by different researchers, sometimes in relation to knowledge and data, sometimes set apart. Claude Shannon, for example, viewed information purely as communications; Buckland saw it as a physical thing; Bates in relation to evolutionary psychology; Hjørland as signals via semiotics (Bawden and Robinson 2012); and Brookes (1980) and Farradane (1979 and 1980) in the late 1970s and early 1980s took opposing views about information. And that is just in the West: Belkin (1975) reported on Soviet views of information and, no doubt, different views exist around the world.

In addition to this, researchers have also approached information from a philosophical point of view while others have used models from areas such Biology and Physics - Luciano Floridi (2010 and 2011) looks at all these areas while proposing a philosophy of information.

Researchers within Library and Information Science have also made approaches to knowledge, such as the two opposing models discussed by Bawden and Robinson (2012): an approach adapted from Karl Popper's Worlds and the D-I-K-W pyramid model proposed by Ackoff.

In an earlier essay for this course (Atkinson 2014), I looked at the work of three LIS writers (Michael Buckland, Marcia Bates and Birgir Hjørland), in seeking to define information. Little consensus was found and this dissertation proposes to expand that essay and examine further the relationship between data, information and knowledge as presented in LIS literature and cross examine these information theories with ideas from other areas within the LIS domain.

A similar approach will be taken in this dissertation as, again, much of what has been written comes from a purely theoretical framework. However, elsewhere in LIS, different models involving information and knowledge may exist separately from Information Theory, for example, in Information Literacy, Information Retrieval and Knowledge Management.

These areas all show interactions between information and knowledge that could be key to creating a more practical understanding of their relationship. It is from this that I propose to introduce the concept of Dust, which will be further explored in the literature review below.

The definitions of data, information and knowledge will be fully explored in this dissertation, as information was already in a previous essay (Atkinson 2014), as will the relationship between these three concepts.

1.1 A Statement of Aims

The aim of the dissertation is:

- to comprehensively review and discuss the Library and Information Science literature on the relationship between data, information and knowledge

The objectives of the dissertation are:

- to try and find consensus and points of agreement, or indeed disagreement, within this area of research;

- to investigate if there is a better way to explain the issues through the study of other Library and Information Science research areas in order to try and find a common ground for consensus;

- to further explore an idea from an earlier essay that Philip Pullman's idea of Dust from the *His Dark Materials* books may form a better way of looking at this issue; and

- following on from the last point, to investigate the author's contention that knowledge should be more involved in information theory, that it both precedes and follows information; and potentially to suggest a new model of information

2. Literature Review

This section will present the findings of a literature review on this topic, divided into Background, Definitions of Data, Information and Knowledge and the Relationships between Data, Information and Knowledge.

2.1 Background: A Short History of this Topic

As reported in the Introduction, meanings to the concepts of Data, Information and Knowledge, and most especially Information, have been sought for decades with the LIS domain. This has been more the case since the Second World War when the term Information Science was coined and academics within the field needed to define themselves; which they have been trying to do ever since.

Buckland (2012a), though, shows that this goes back to the nineteenth century as people like Paul Otlet, Suzanne Briet and Melvil Dewey used spatial analogies for knowledge organisation - Otlet in particular spent many years trying to organise all the information in the world (van den Heuvel and Boyd 2011). Others such as Emmanuel Goldberg and Vannevar Bush sought other physical ways to organise information and knowledge (Buckland 1992). However, it was the work of Claude Shannon (1949/1969) that started to bring such work more into the theoretical sphere, through his Mathematical Theory of Communication (MTC).

Shannon's Theory, expressed as an equation and diagrammatically, has been very influential on various fields (Gleick 2011) and has caused much discussion within LIS. Bar-Hillel and Carnap brought semantics into information, pointing out that the meaning mattered (Cornelius 2002, Machlup 1983). Since then, ideas about information have grown, with various writers adding in data and knowledge to the fore, with Jesse Shera and Birgir Hjørland being noticeable adherents to bringing social epistemology into LIS.

At different points in time, various writers have added to the LIS literature in this area- mostly writing about information foremost, but often writing about data and knowledge at the same time. At the launch of the *Journal of Information Science*, Bertram Brookes and Jason Farradane each offered their own view of information and its relationship to knowledge; an area Neil Belkin was writing about at about the same time.

Either side of 1990, Michael Buckland (1991) proposed that Information was "a Thing" and Ackoff (1989) put forward his influential Data-Information-Knowledge-Wisdom (DIKW), which has been altered and critiqued by various writers (such as Frické 2009 and Ma 2012). While in the first decade of the 21st Century, Marcia Bates and Birgir Hjørland debated their views on the nature of information and by the end of the decade, Floridi had developed his Philosophy of Information.

This is just the tip of the iceberg- Bates (2010), Bawden and Robinson (2012), Cornelius (1999), Capurro and Hjørland (2000) and Machlup (1983) are just a few of those to have written reviews of the topic. While Zins (2007a) and Hartel (2014) each have performed interesting and very different surveys to add to the discussion; and below will also be discussed further views and critiques of the LIS approach to the relationship of data, information and knowledge.

All of which will be looked at in more detail below and discussed in Section 3. First, however, I must present my Methodology.

2.2 Methodology

In order to find articles to assist in the research of this topic, I ran searches in the Library, Information Science & Technology Abstracts (LISTA) database and on CityLibrary Search. I also extensively used existing reviews on the topic. The latter was to ensure I covered the main writers on the topic while the former meant I would include anything newer in the literature than those reviews and also pick up more of the later critiques covered.

With LISTA I set up a series of searches covering aspects of this topic having come up with different series' of related terms that I could combine with the Boolean operator OR (data OR information OR knowledge OR "information theory" theory; library studies OR library science OR information science OR (library and information science) OR LIS; data OR information OR knowledge OR "information theory"), each one searching the abstracts of articles. I then combined these searches using the Boolean term AND in order to refine my initial results.

In City University London's discovery tool, CityLibrary Search, I used subject searches to help ensure I had good coverage of the topic. Additionally, while searching for individual articles to read via CityLibrary Search, other results would pique my interest and become involved in the study - Belkin (1975) and Buckland (2012a) being notable examples.

Finally, the LIS literature itself was used to help get an initial grasp on the topic. Bawden and Robinson (2012) and Bates (2010) were foremost in this, each giving comprehensive reviews on the topic that could be used to build a bibliography. From reading articles cited, I was introduced to further articles cited, allowing a bigger picture to be built. Additionally, I had, of course, already written on a small section of this topic and revisited those articles cited in Atkinson 2014.

For the most part I was able to gain the books and articles themselves through City University Library's resources, either in print or online, but where I could not I used their Inter-Library Loans service.

2.3 Definitions of Data, Information and Knowledge

2.3.1 Introduction

Articles written about the definition of information often start by exclaiming the amount of definitions that exist on the subject, one even going so far as to call it a cliché to even mention it (Furner 2015). However, as far back as 1973, Whittemore and Yovits pointed out that information was one of the most overused words in the English Language.

Alongside this are a plethora of definitions for data and knowledge too, as well as different models illustrating how the three interact.

The following section will examine each of these areas in turn so that they may be discussed in the following chapter.

Before assessing the relationship between these three concepts, it is important to look first at the definitions that have been offered as this has a knock-on effect.

2.3.2 Data

"Data, data everywhere and not a thought to think." - Jesse Shera (Machlup 1983)

Generally in the LIS literature, data is discussed less than information and knowledge, the latter two in the triangle becoming the dominant pairing with data often becoming

synonymous with information. Cornelius (2002) goes as far as to state that data does not interest him.

As words (in English), data is the youngest of the three (Bawden 2001), originating in the 16th Century and being given its modern meaning around 1950 with the advent of the computer. The Google Books Ngram Viewer (see Figure 1) shows indeed that the use of the word data increased rapidly during the 20th century, particularly in the latter half. Bawden (2001) expands his look at data by including different systems analysis views, reporting that "hard systems proponents" see data and information as synonymous, while "soft" proponents see data as information that lacks meaning.

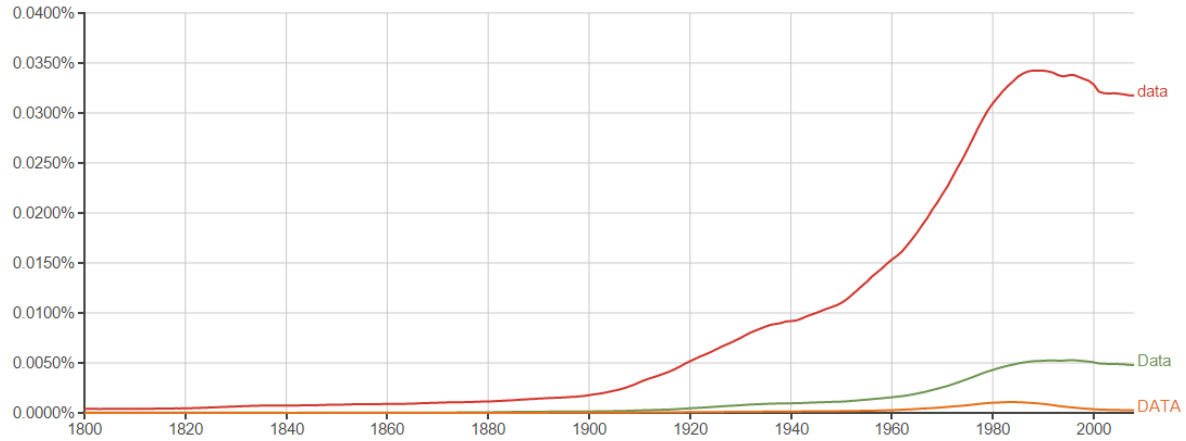

Figure 1: Google Books Ngram Viewer on Data

Machlup (1983) bemoans the uses of all three of the terms data, information and knowledge, calling for the dictionary definitions to win out. As such, he points out that Data comes from the Latin word datum which means "the given" and so sees data as the things that have been given to the "analyst, investigator, problem-solver," no matter what form they are in. Machlup also concedes that data can also be seen as anything fed into a computer. In this, Machlup possibly picks up on the difference of what data is to machines and humans (Ma 2012), reflecting the shift in meaning that Bawden (2001) said happened in the late 20th Century.

Two approaches can also be seen in Bates (2005) who presents two definitions of data. One of these is data 1: the information perceived by an animal and made into information 2. In other words, what is given to a creature, rather than that which is sought, which forms Bates' second data definition, data 2: information selected for social purposes, i.e. data generated by the researcher. As suggested, Bates sees two types of data here- that which occurs naturally or is discovered, and that which is purposefully discovered or sought. In both cases data is either synonymous with information or that which will be converted into it.

That data is something that will be processed to become information is a common definition of data within LIS. Buckland (1991), for example, relates data to his definition of information as thing: "The term "information" is also used attributively for objects such as data and documents." Although Buckland is not massively explicit in giving a definition of data, other than its Latin origins, as seen in Figure 2 below, Buckland saw data as a tangible thing, synonymous with or similar to a document, and making up his definition of Information-as-thing.

	INTANGIBLE	TANGIBLE
ENTITY 2.	Information-as-knowledge Knowledge	3. Information-as-thing Data, document
PROCESS 1.	Information-as-process Becoming informed	4. Information processing Data processing

Figure 2: Buckland (1991)'s Four aspects of information

Similarly to Ackoff (1989), data "are symbols that represent properties of objects, events and their environments. They are products of *observation*." (italics in original). He goes on to liken data to metallic ores which are of no use until processed (by management information systems) into the desired product.

Ultimately data is seen by Ackoff, and others, as the rawest form of information, as is perhaps shown best in the Data-Information-Knowledge-Wisdom (DIKW) pyramid (see Figure 12), or at the start in the linear form Data - Information - Knowledge - Wisdom. To Ackoff, writing about management information systems, data is the start of a process that ends with understanding and wisdom helping to improve efficiency. Like others (Buckland 1991, Bates 2005), Ackoff sees data as nothing until it has been processed.

In discussing DIKW, Frické (2009), having presented Ackoff's definition of data, agrees data is of no value until it has been transformed into a relevant form. In this respect we see that data is often defined as information lacking processing. As Frické notes, data can be gathered by machines but is not useful until analysed or interrogated by an enquiry. Another example is Floridi (2010), who calls data unmeaningful information.

Checkland and Holwell, in reviewing the DIKW model, add capta to data as an added concept (Bawden 2001). Capta is defined as data that you are interested in, adding also to Machlup's definition of data and perhaps coming into line with that of Bates' Data 2. From all the data a problem-solver or organism is given, the capta is what they choose to use.

Looking ahead, it is possible that the approach to data could change again. In talking about data and information retrieval systems, Furner (2015) talks about "research into systems for collecting, organizing, and analysing the "big data" produced by unprecedentedly large-scale projects." He adds that there are three sorts of data sets: "1) the universal set of all kinds of data, structured and unstructured; 2) structured data; and 3) unstructured data." Here, there are two issues that have not been covered above.

One is that "structured data" within information retrieval systems - the data such systems search, which come from cataloguing and subject heading schemes. In other words, metadata. As defined above, data is contained within books, found within nature or gathered for a specific purpose. Metadata, rather, is created as an intellectual tool to be stored and used to serve a function: to help users find documents.

The other is the "structured data" in the growing storage of very large data sets for analysis. A good example of these are those that have been collected and stored by City University London lecturer, Ernesto Priego on City Research Online. These are mostly (there are 23 at the time of writing) archives of Tweets made during and around conferences and data generated by Altmetrics of articles on particular topics.

As documents, these could be seen as information but they are available for use as raw data for people to analyse as they would like. As digital storage becomes cheaper and journals move online, there is the possibility for data generated through research to be posted alongside articles discussing the results so that others can analyse further and draw their own conclusions.[1]

At City University London, at least, Priego is very much a pioneer as his are the only datasets on City Research Online. However, it is clear that such collection and dissemination of data could, in the future, have a greater bearing on the LIS definition of data. Certainly, as more datasets are published online, it may become easier to agree with Jesse Shera's quote from the start of this section.

Within the LIS literature, though, data is generally conceived to be information before it becomes that. It can be tangible, as Buckland (1991) and Machlup (1983) suggest, or environmental as Bates (2005) suggests with Data 1. Bates and Checkland and Holwell both add an extra definition to separate out all data from that which is selected to be used, or given meaning.

2.3.3 Information

"Information mean news or facts about something" - Sesame Street's Cookie Monster (Losee 1997 and MarshalGrover 2008)

"Information is not under-theorized" - Cornelius (2002)

2.3.3.1 Introduction

As Ian Cornelius suggests, information is a much more talked about concept in LIS than data or knowledge due to its nature as a titular concept in the field. Many definitions have been put forward to try to explain the "information phenomenon" which have an impact on the relationship between data, information and knowledge.

To some extent the definition of information has already been discussed (Atkinson 2014). There, it was shown that "from looking at the etymology and definitions of the word," Capurro and Hjørland (2003) and Bawden (2001) show that even trying to use a dictionary to solve the issue is problematic" and proceeded to discuss the definitions of information presented by LIS writers, Michael Buckland, Marcia Bates and Birgir Hjørland. Please see Appendix A for this essay, elements from which will be used for further discussion below.

Three articles in particular have neatly shown the differences in the opinions of what information is. Zins (2007a) tasked academics across the world to define data, information and knowledge with a great variety of results. Bates (2010) in the *Encyclopedia of Library and Information Sciences*, gives a thorough rundown of different approaches to the term information, and many of the writers and articles referenced will be looked at in more detail here. Hartel (2014) took a similar yet different approach to Zins (2007a) by asking new Library School students in Toronto to draw what they thought information was in a small square and describe it in words on the back.

All three articles offer various definitions of information with Bates and Zins offering overviews of opinions from within the LIS sector (either over time or at the time of writing),

[1] This idea came out of a discussion in a INM356: Research, Evaluation and Communication Skills session led by Andrew Robson on evaluating research in which the class felt that presenting research results in such a way would be a good step forward in transparency.

while Hartel's study gives a more formative insight. Bates concludes that the "core concept of "information"" in IS is "highly contested" and is seen in the following ways as "a proposition, a structure, a message, or an event," that it can contain truth or be indifferent, it, "can be socially embedded, under perpetual re-interpretation, or be measurable" or is, in fact, "a worn-out idea deserving of dispatch."

Similarly, the study of Zins (2007a) shows polar views in the approaches academics take to information and, in this case, its relationship with data and information. Among these are Metaphysical vs Nonmetaphysical Approaches and Human Exclusive Versus Nonexclusive Approaches.

The pictures created on the "iSquares" used by Hartel (2014) again show a plethora of views about information, from the human to the physical (books), machine (binary information, computer networks), and the abstract (ordered patterns).

Information, clearly, can mean all sorts of things to different people. A brief history of information definitions will show us in more detail these differences in opinion - and it is one littered with differing and opposite views.

2.3.3.2 The Beginnings of Information Theory

Following the Second World War and while working at Bell Laboratories, Claude Shannon developed a theory of information that has become influential in many fields (Gleick 2012), including LIS.

For Shannon, information is related to messages moving along a telephone line and his concern was with the accuracy with which information would be passed along it. In this respect, Shannon saw information as a statistical probability of receiving the correct message (Machlup 1983).

Shannon saw only a practical application in information, the meaning of the message was irrelevant and, though his ideas have been influential in LIS, many people have held quite a different view of information. Weaver, who wrote an essay on Shannon's work, and published the two alongside one another, transplanted human communication onto Shannon's work as instructions (becoming the first person to adapt Shannon's theory). Even so, Machlup (1983) still found that Shannon-Weaver's theories as significant only for the design of communication systems.

At about the same time, Wiener, who was working in the field of cybernetics held an almost opposite view of information to Shannon. As Cornelius (2002) reports, Wiener saw information as negative entropy and a degree of organisation whereas Shannon had seen it as the reduction of uncertainty and entropy itself. And where Shannon saw information within (telephone) communications, Wiener saw it as something higher than human communication: "Information is information, not matter or energy." (Capurro and Hjørland 2003).

2.3.3.3 Brookes vs Farradane and Bates vs Hjørland

In the late 1970s and early 1980s, Bertie Brookes and Jason Farradane each wrote about information in the early issues of the *Journal of Information Science*, offering different views from one another (Bawden 2008): Farradane seeing information as physical, and Brookes saying it is not.

Farradane (1979) states that information is a physical thing, a surrogate of knowledge; i.e., that information is to be found in books, journals and other physical manifestations of

people's thoughts and work. He goes on to say that information is the only part of the communication chain that is physical. Brookes, however, disagreed, seeing information as "fragmented knowledge" (Brookes 1981 via Todd 1999 and Bawden 2008) and that it existed in the cognitive realm along with knowledge.

This is quite a fundamental difference and one that has a knock-on effect on how people could view or approach LIS and what it should be dealing with. Is information entirely held within the mind, and non-physical, as part of a communication process; or is information physical and held in documents?

As noted in Atkinson (2014), Bates and Hjørland debated the nature of information and ultimately came to a similar difference of opinion as between Brookes and Farradane. This is, that they each see information as different points of the same process. As was pointed out then, Bates sees information as both subjective and objective, while Hjørland as subjective only and "as the completion of a process that results in someone becoming informed" (Atkinson 2014). Bates agrees with this but stretches information much further, calling it the "the pattern of organisation of matter and energy". Additionally, where Hjørland sees information as a human phenomenon, Bates applies it to all creatures.

The difference, though, is based in semiotics and in this difference we hit upon the argument of whether information is a sign or a signal; whether it can be both objective and subjective or is purely subjective. Each view will have an effect on information's relationship with knowledge and will be discussed further below.

2.3.3.5 Other views of information: Floridi, Losee, Furner

Philosopher of information Luciano Floridi, has developed his own General Definition of Information (Floridi 2010 and 2011), which he views as semantic, saying information must be "well formed, meaningful and truthful data." As Bawden and Robinson (2012) point out, this means true information must be separated out from mis-information, pseudo-information and false information; and that Shannon's view of information is seen a structured data.

Floridi has produced a map (Figure 3, below) to show the flow of data through to knowledge, and uses examples from maths, physics, biology, engineering and economics to explain each step of the process (Floridi 2010). Meaningful and relevant information then becomes knowledge.

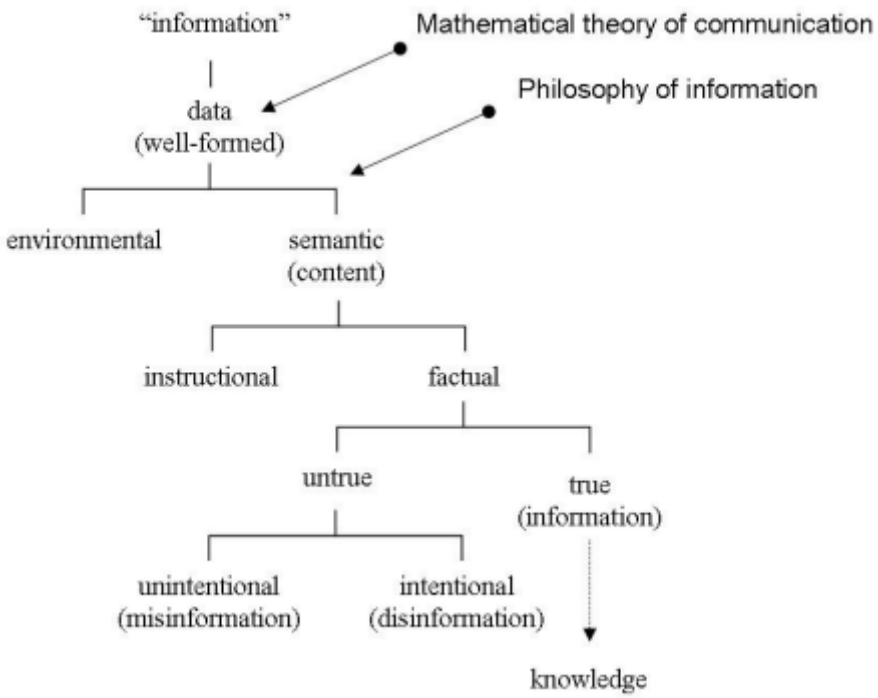

Figure 3: Floridi's Information Map (Floridi 2009)

Robert Losee (1997) sought to find a discipline free definition of information, ultimately saying that information "may be understood as the value attached or instantiated to a characteristic or variable returned by a function or produced by a process." This, he believes, allows for different ideas about information (as communication, as thing or as knowledge) and for ideas of information as a sign or signal (as "the value returned by a function is informative about the input to the process and about the process itself.") The idea is that anyone from any discipline, and those who disagree about what discipline LIS is, can use the same terminology and data when talking about information.

Bates (2010) questions Losee's work, which says information is in the value at the end of a process by asking "What about situations that are not processes?" She also argues that his example of baking a cake, which he says "provides information about both the process and the original ingredients," clearly does not. With a cake, says Bates, the end result can be misleading and someone who has never seen a cake would not be able to tell you anything about the process or ingredients.

2.3.4 Knowledge

"What is knowledge? The question sounds simple, but the answer is difficult." - Qvortrup (2006)

"Full knowledge is better than half-knowledge." - Iorek Byrnison in The Amber Spyglass *(Pullman 2000/2011)*

2.3.4.1 Introduction

The word knowledge originated in around 1300 in Northern England, before appearing in the south in the late 1300s, its most common meaning being the "antithesis of ignorance" (Bawden 2001). As Nagal's (2014) book on Knowledge shows, this has been a topic in philosophy since the days of Ancient Greece with every philosophy movement in history becoming involved.

Meanings within the LIS literature vary a little with some arguments over its use. Indeed, as noted by Bawden and Robinson (2012), views within LIS often revolve around two "useful [but] incompatible" models- those of Popper's Worlds and Data-Information-Knowledge-Wisdom (or DIKW).

In Popper's Three Worlds, World 1 contains documents and physical information, World 2 subjective knowledge and World 3 objective knowledge, the three interacting together (Bawden and Robinson 2012).

DIKW, however, has a different view of knowledge and its relationship with data and information. Rather than interacting worlds, the concepts are placed in a linear progression or pyramid, leading one from data to knowledge and beyond. Here, knowledge is a form of "refined information" and it can be either external or internal (Ackoff 1989).

Beyond this, views of knowledge in LIS often take their cues from these two ideas, though sometimes placing knowledge on any one of Popper's Worlds, or only 2 or 3. Others look to social epistemology, seeing knowledge as shared across society and subject domains, rather than simply being one individual's 'property.'

2.3.4.2 Knowledge in LIS

Bates (2005) sees knowledge as information given meaning and integrated with other contents of understanding. To her, knowledge is something entirely within our heads and dies with us (Bates 2006). She also believes truth is not a requirement of knowledge, as it is a kind of meaningful belief, as opposed to the justified true belief of Ancient Greek philosophy.

Additionally, there are views of knowledge in LIS that cast it across society and not just about an individual. Jesse Shera's social epistemology, for example, as reported by Furner (2002), looks at how "society as a whole seeks to achieve a perceptive or understanding relation to the total environment." Shera saw epistemological knowledge as potentially the grounding theoretical basis for LIS, "the analysis of the production, distribution, utilisation of intellectual products." Shera said "graphic records" contain "recorded knowledge" (Furner 2002), meaning that knowledge resides in the mind but can be recorded in physical documents, and this is subjective knowledge.

Jesse Shera (and Margaret Egan, whom Shera initially worked with before her death) saw two types of knowledge: "*intrinsic* knowledge of the kind that the individual has, and *extrinsic* knowledge of the kind that society has" (Furner 2002, italics from original). The biggest difference, thought Shera, was that individuals have a more emotional experience with knowledge, "and [are] thus capable of understanding in a way that society is not" (Furner 2002).

Shera saw social epistemology as “the solution to librarianship’s search for an intellectual foundation” (Bawden and Robinson 2012), that a librarian would need “a true mastery over the means of access to recorded knowledge” through both understanding that knowledge and its role in the librarian’s within society (Bawden and Robinson 2012).

Capurro and Hjørland (2003) and Hjørland (2007) continued this view via a domain analysis approach. They see information as subjective but not primarily in an individual sense as this information is also shared across subject domains. For example, different information might be seen in a rock by people from different disciplines. In other words, knowledge brought to the table, “extrinsic knowledge” as Shera would have called it, will affect what you come away with. Hjørland (2002) also calls this pre-understanding, suggesting that knowledge comes before information.

Zins (2006), in seeking to redefine information science as knowledge science, casts the ideas of knowledge as being a thought versus a thing, i.e. knowledge being encased in someone’s mind versus knowledge in documents. Zins sees this opposition as two domains: the subjective domain and the universal domain. In the subjective domain, data, information and knowledge are all related to sensory perceptions, placing knowledge in the individual’s mind. Whereas in the universal domain, data, information and knowledge are physical, meaning knowledge is a set of symbols representing the meaning or content of thoughts that are justifiably believed to be true.

In criticising definitions of the knowledge society, Lars Qvortrup (2006) found there was no “sociologically relevant” definition of knowledge. In answering the question, “What is knowledge?” he uses the view from his window while writing, of snow melting from fir trees, and wonders how he knows what these things are, when the things themselves have no awareness. From this Qvortrup, shows that knowledge is something humans have created and thus, knowledge is “confirmed observations” that can be confirmed over time or by society. When reconfirmed by the self, an observation becomes personal knowledge and when it is confirmed by another it becomes social knowledge.

As in Hjørland’s domains, Qvortrup says that knowledge can change between societies or organisations who might make different observations- much like Hjørland (2007; see also Atkinson 2014 for further discussion about Hjørland) saying that a geologist and an archaeologist might interpret the same information from a stone differently.

From such observations, Qvortrup builds up four orders of knowledge from simple facts, to knowledge about knowledge, knowledge about the criteria of knowledge and knowledge about the boundaries of possible knowledge. These categories are presented here in Figure 4:

Knowledge category	**Knowledge form**	**Knowledge designation**
1st order or simple knowledge	Knowledge about something	Factual knowledge
2nd order or complex knowledge	Knowledge about the conditions of knowing	Reflexive or situative knowledge
3rd order or hypercomplex	Knowledge about the conditions of the reflexive knowledge system	Systemic or creative knowledge

knowledge		
4th order knowledge	Society as dynamic knowledge horizon, i.e. the knowing society	World knowledge

Figure 4: Qvortrup's Orders of Knowledge (adapted from Hjørland and Nicolaisen 2005/2007)

Finally, an interesting point that comes out of LIS is the rejection of knowledge as justified true belief, as set out in classical philosophy where Plato maintains that information must be truthful in order for it to become knowledge (Nagel 2014).

Although Floridi (2010) and Ackoff (1989) both see truth as being necessary to information and knowledge, Frické (2009), Bates (2005), Zins (2006), Buckland (2012b) and Machlup (1983) all hold different views. A fuller explanation will be given in Section 3.2.2.

2.3.4.3 A Note on Understanding

"This is a different kind of knowing… It's like understanding, I suppose." Lyra Belacqua in Northern Lights *(Pullman 1995/2011)*

In Ackoff's (1989) original take on DIKW, he places understanding (an "appreciation of why") between knowledge and wisdom, something that later writers have often left out (Bawden 2001). Jonathan Kvanvig (2003) says that some argue the word knowledge should in fact have been translated from its Ancient Greek root as understanding and wonders if knowledge, in our information age, has been boiled down to mean only the sort of facts important to be good at *Trivial Pursuit.*

For Kvanvig, 'Understanding' is separate from information, knowledge and truth: "understanding requires the grasping of explanatory and other coherence-making relationships in a large and comprehensive body of information. One can know many unrelated pieces of information, but understanding is achieved only when informational items are pieced together" (Bawden 2012a).

For Kvanvig, Bawden (2012a) continues: "The object of understanding that which is understood, is, for Kvanvig, not a number of single propositions… but rather an "informational chunk". The mind then has to deal with chunk, and appreciate it, dealing with any "ambiguity, contradiction, missing information," or any other issues, and needs to make sense of these before information can become knowledge.

Bawden (2015) and Bawden and Robinson (2015a) have recently made further arguments for understanding via a conference paper and subsequent journal article. Each reviews the work on understanding from LIS and concludes that more needs to be done.

2.3.5 Summary: The Survey of Chaim Zins

As can be seen from the above, there is no real consensus within LIS on the definitions of data, information and knowledge, although there are many similarities. This has perhaps been shown best by the survey taken by Chaim Zins (2007a) which asked LIS academics from around the world for their definitions of data, information and knowledge.

A read through shows all kinds of definitions coming up, some of which have been covered already above. Generally, data is seen as symbols, that either within or without a computer system, that are yet to be used or assigned meaning; information is this data having been

processed, though this could be physical or in the mind; while knowledge is often seen as information, received and understood.

As with Zins's (2007a) discussion, although this gives a flavour of the discord within LIS, what we are really interested in here is the relationships between data, information and knowledge.

2.4 The Relationships Between Data, Information and Knowledge

"Putting the three concepts ("data," "information," and "knowledge") as done here, gives the impression of a logical hierarchy… This is a fairytale." Rafael Capurro in Zins (2007a)

2.4.1 Introduction: Paul Otlet's Spatial Analogies

These definitions lead onto a variety of differing opinions on how data, information and knowledge interact with one another along the flow from source to brain.

Figure 5: Paul Otlet's *Index scientiae* spatial analogy (Buckland 2012a)

A 2012 article by Michael Buckland (2012a) looking at spatial analogies relating to knowledge organisation showed various ways Paul Otlet, the early 20th century documentalist, and others, described or showed knowledge in analogies.

One, pictured here (Figure 5), shows a fruit-bearing tree rising from a set of books, a key upon a rope is wound around the tree trunk and the term *Index Scientiae* (Index to Knowledge) hovers in mid-air.

This itself gives an interesting take on the relationship, if not between all three concepts, then at least that between information and knowledge as knowledge, seen in the fruit, grows from information in the books which replace the ground as a fertile soil for imagination. The presence of the key and the term *Index Scientiae* show the process one must go through to gain the (new) knowledge, i.e. a key can unlock doors just as an index allows access to recorded knowledge.

As Buckland interprets this, Otlet is really talking about old knowledge from the books becoming new knowledge in the fruit but the books could easily be interpreted as information or data, just as the key to access have now expanded from indices to databases and search engines. The key could also be seen as understanding.

What is certain, however, is that Otlet, in this picture (and others discussed by Buckland 2012a), is showing the relationship between books and knowledge and, thus, this is an early representation of the relationship between data, information and knowledge in LIS.

2.4.2 Shannon-Weaver: Criticisms, Praise and Partially Discounting (an) Information Theory

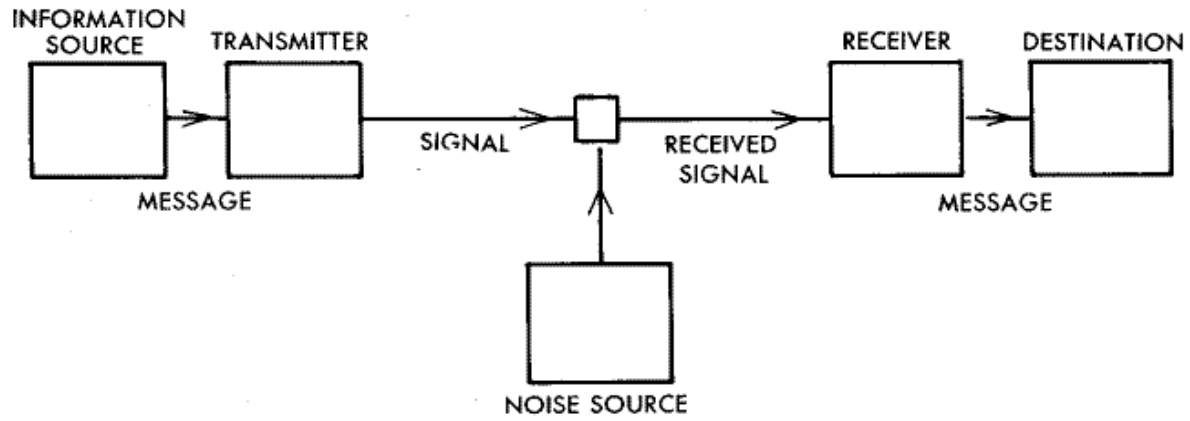

Figure 6: A Representation of Shannon's Ideas (Shannon and Weaver 1949/1969)

In presenting the flow of information from source to receiver along a telephone line (see Figure 6 above) and working out the probability that the correct message would be received (represented by the equation $H = -\Sigma\, p_{(i)} \log p_{(i)}$) (Shannon and Weaver 1949/1969); what became known as information theory was born. From this a lot of information theory in LIS has grown, although much has also grown away from this starting point, using it only as that.

Given that information is often discussed in relation to data and knowledge within LIS literature, it could be argued that Shannon's "information theory" should give an insight into this relationship.

However, Shannon was interested only in the transfer of information and what affected its correct transmission. He saw information as a function of freedom in selecting a message (Bates 2005). This was a very practical approach, really; as Shannon was a telephone engineer, he was interested in ensuring a call was heard correctly at the far end- that the message was received without being destroyed by noise on the line, and, as such, created an equation concerning this using entropy as a word to describe the chaos to be avoided.

This idea certainly seems to have been seen as useful by some LIS writers since its original publication. Brookes (1980) talks of the possible application of Shannon's measurement of objective information as potentially being applied to objective knowledge. Here, though, Brookes also makes reference to Shannon's use in telecommunications systems and computers, perhaps inadvertently pointing out its limitations in LIS. Hjørland (1998), however, explicitly says Shannon's ideas are a part of computer science.

As various LIS writers have pointed out, this is the essential limitation. Ellen Bonnevie (2001) talks of the rejection by Machlup of Shannon's Mathematical Theory of

Communication (MTC) saying that information is addressed to human minds and is received by them, and that information as a description of signal transmission is a poor use of the word. As with Hjørland, Machlup (1983) saw MTC as of significance only for the design of communication systems, pointing out that Boulding and others dismissed it as irrelevant for the social sciences.

Ma (2012) continued this argument by pointing out that what Shannon was talking about was an engineering issue onto which Weaver transplanted human communication; but, crucially, without approaching semantic issues at all. Weaver only looks at procedures, cutting out all understanding and meaning, and sees success as the carrying out of instructions correctly once the message has been received. Ma calls this a "distorted picture of human communication," leaving out, as it does, all cause and effect and negotiated meanings (as will be seen in other models below).

Bawden (2001) calls the use of the word information by Shannon "unfortunate" saying that "'signalling theory' would have been more appropriate." He goes on to point out, though, that Shannon's methods have been put to good use in information physics and biology to connect information to the wider universe and the human genome project.

Arguably, however, Shannon's model can quite easily be taken and applied to any number of problems, and was picked up quickly within Communication Studies (Gleick 2011). This is something that will be returned to below in consideration of how Shannon could be applied to the relationship between data, information and knowledge within another area of LIS.

2.4.3 What LIS Writers Say

While writing about "the shifting terminologies of information" in 2001, Bawden (2001) notes that as there are no agreed definitions for the terms data, information and knowledge, there are equally no agreed set of relations between the terms. Even the DIKW model has been expressed in different ways (Rowley 2007, see Figure 12): such linear progressions have been called "a fairytale" by Capurro (Zins 2007).

Below, however, is a summary of different ideas from within LIS about how data, information and knowledge relate to one another which will be discussed in the following section.

2.4.3.1 Brookes and his Fundamental Equation of Information Science

In 1980, Bertram Brookes published the final version of his fundamental equation for Information Science, which he had been working on and refining via a series of articles published through the 1970s (Todd 1999). A year later Brookes called the equation a skeleton, something for future information scientists to build upon and flesh out (Todd 1999).

The equation, $K[S] + \Delta I = K[S+ \Delta S]$, shows that a knowledge structure ($K[S\}$), when given a new piece of information (ΔI) is changed to a new state ($K[S+ \Delta S]$) consisting of the original knowledge plus the changes put into place by the new information.

For Brookes this was key. His 1981 paper *Information Science and the Science of Information*, concludes by saying that the information concept is empty: what should be important to LIS is the relationship between information and knowledge. This relationship, aside from the equation, Brookes defined as follows: "knowledge is structured integrated information and information is fragmented knowledge." To Brookes, there was an intimate relationship between the two, so much so that they are almost synonymous.

Todd (1999) and Bawden (2011) each discuss this equation further in useful ways. Todd explains that the equation shows that what goes in depends on what is already there, that knowledge is not just about the input of information but the interaction of it with the current knowledge state to create the new one.

Todd also talks about how Brookes saw a practical and theoretical side to Information Science. In talking about the latter, Todd explains that Brookes felt this theoretical side was about "the interaction between the private, inaccessible thoughts and mental images of people, each unique; and the public documented artefacts of knowledge" (Todd 1999).

Which is, perhaps, another way to say the interaction between information and knowledge. To combine the two ways of saying this, Brookes saw knowledge as a structure entity changed evermore over time by information.

Furthermore, Todd reports on articles by Brookes in which he goes further into this interaction by explaining how physical information is converted to neural energy for transmission to the brain where it interacts with older knowledge to create the new. This Information-Knowledge relationship was summarised by Todd (1999) in Figure 7:

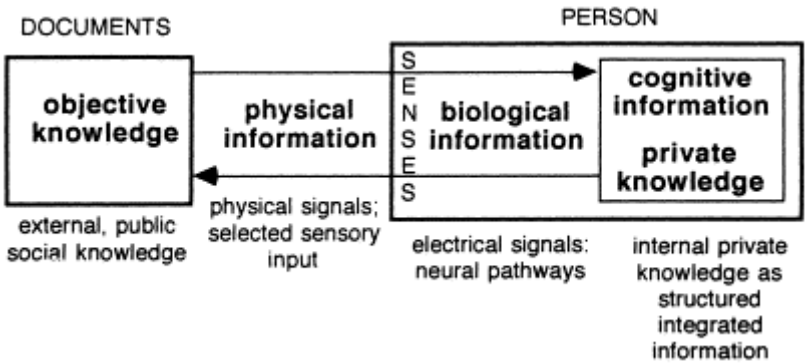

Figure 7: Brookes' Information-Knowledge Relationship from Todd (1999)

Here we can plainly see the full extent of Brookes' ideas, including the different types of knowledge and information he proposed. In the world outside the mind is the pool of objective knowledge contained in documents and society. This can be turned into physical information in the form of signals that are converted by the senses to biological information that is processed by the brain into the cognitive information that changes one's private knowledge. The reverse process then allows private knowledge to become objective knowledge.

Bawden (2011) raises some other interesting points relating to Brookes' equation. Firstly that ΔS is "an effect of the modification" caused by the addition of information and that the "same information may have different effects on different knowledge structures." Therefore we see that the knowledge created by new information is affected by the knowledge already present. As Brookes (1974) put it, "Crudely interpreted, it says that whatever 'goes' in depends on what is already 'there'."

Another way in which Brookes looked at this relationship was in conjunction with the ideas of Karl Popper's Worlds. Brookes (1980) discusses his own view of Popper's Worlds, which he feels lend themselves very nicely to LIS. In particular that "World 3 should commend itself to library and information scientists because, for the first time, it offers a rationale for their professional activities which can be expressed in other than purely practical terms."

These “practical terms” he saw as being the role of looking after all that different disciplines deposit there from their activities on the other Worlds, for example: “Natural scientists and technologists explore and exploit World 1 and deposit their records and artefacts in World 3. Social scientists and humanists study and reflect upon World 2 and the interactions of World 2 with World 1; they too deposit their records and artefacts in World 3.” Again, the theoretical side he saw as organising this knowledge, rather than documents.

Various authors have criticised Brookes’ equation. Ian Cornelius (2002) raises a point later mentioned also by Bawden (2011) that the equation suffers from a lack of measurement, though, as Bawden (2011) says, Brookes himself did not see this as a mathematical equation. Cornelius, however, goes on to say that designers of systems would need this element to make use of it and that such a lack of structure gives the equation a “water into wine” effect when talking of turning information to knowledge. Cornelius develops the latter point into saying that Brookes, really, is talking about understanding. This is backed up a bit by Brookes (1974) himself who, in explaining an early form of his equation says that it sums up the process of learning.

Cornelius (2002) additionally raises problems with Brookes’ use of Popper’s Worlds, talking about the “seeming conflation of information as a part of knowledge” is problematic in making a theory of information.

Ingwersen approached issues with Brookes’ equation by adding to it at the start and end to account for potential information. In his version, Ingwersen inputs potential information at the start which is then mediated by knowledge and produces an end product of potential information. This allows for the production of documents from documents, or the retrieval of information from a computer system. This is because Ingwersen says Brookes does not allow for subjectivity, only an individual’s perceptions and states of knowledge (Cornelius 2002).

Ma (2012) critically reviews Brookes’ use of Popper’s World 3 to form his fundamental equation. In particular, she says that Brookes cuts out human minds in saying that only “human thoughts embodied in artefacts” can alter the knowledge structure. She goes on to say that this reduction of cognition to “an equation and the causal relationship between information and mental states blinds us to how humans really *learn* by reading books, listening to music, and other “mental” or intellectual activities… learning is an activity, just as being informed and knowing are.” (emphasis in original).

In this statement, Ma says that Brookes simplifies the relationship between information and knowledge, suggesting that the act of learning needs to be brought in more to make the equation more realistic; which is in slight opposition to Cornelius’ (1999) statement about understanding and shows how much Brookes’ thoughts may have altered between his statement about learning in 1974 and the article Ma is critiquing from 1980.

2.4.3.2 Jason Farradane

As was stated above, writing at the same time as Brookes was his Canadian colleague, Jason Farradane, who published his own view of information and knowledge alongside Brookes’ in the early issues of the *Journal of Information Science*. Farradane saw information quite differently to Brookes (Bawden 2008), as physical rather than existing in cognitive space.

Farradane (1979), says that information is a physical surrogate of knowledge and that it is the only part of the communication chain that is, or can be, physical (it is “read by,

understood by"). Knowledge, however, is a memorable record of a process in the brain and is only available in the brain (Farradane 1980).

In discussing the transfer of information to the mental state, Farradane (1979) makes some very interesting points. As defined, Farradane says that information is sterile until related to people who are affected by it, that there is no fixed meaning of a document: the only valid meaning is in the originator's thought. As such, language can prove to be a barrier as thoughts can be distorted before spoken (especially impromptu remarks in meetings, Farradane uses as an example) and written language may not be expressed adequately. Getting at the original intention through questioning, suggests Farradane, could itself be difficult as the interviewees' mind structure would alter when questioned.

Farradane also considers the receiving of information by the recipient and the workings of short and long term memory. In doing so he considers the things that could happen when someone receives unrequested information.

The recipient might reject it at the interpretive stage. It could be rejected due to bias against the originator or only accepted for a short period if thought transitory, trivial, or unrelated to larger concerns.

If accepted to long term memory, the information may 1) be added to existing knowledge without changing it, 2) complete a gap in the recipient's knowledge structure or 3) transform an existing knowledge structure if it contradicts what was already known. Reactions to the information could then range from a simple expression in speech to actions undertaken by the recipient.

A year later, Farradane (1980) expanded on these ideas and created a diagram (Figure 8, below) to show the flow of information between an originator and a recipient of that information:

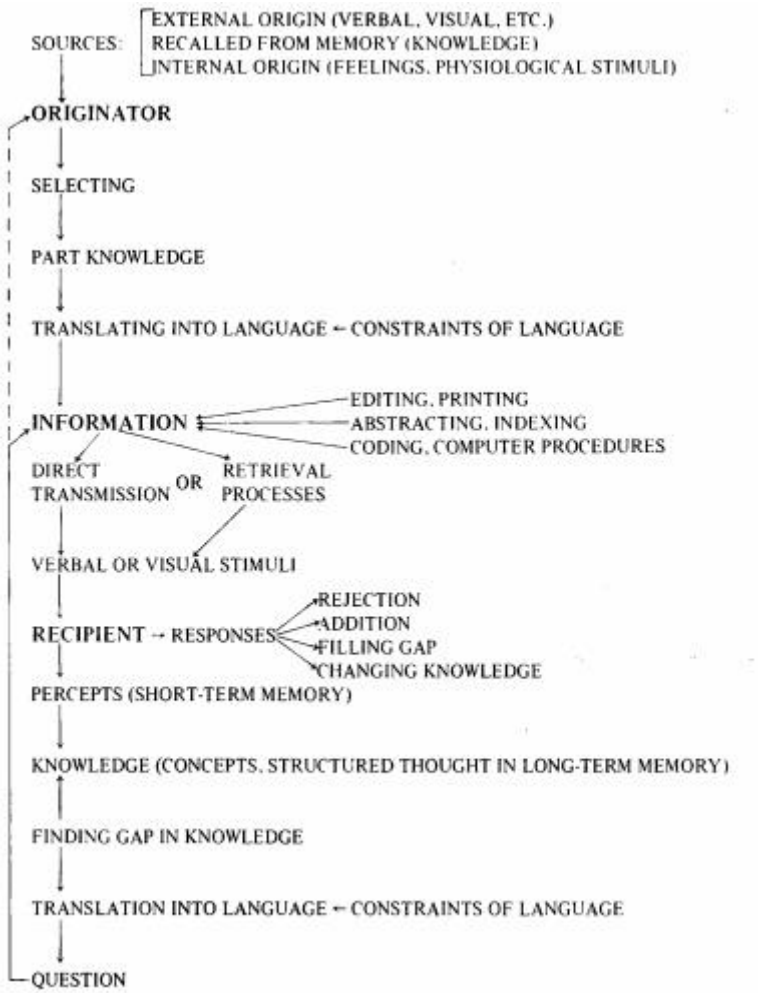

Figure 8: Information passing between a sender and a receiver (Farradane 1980)

In the above, we can see that an originator calls on various sources to create information. These sources include their own pre-existing knowledge and the information is created by selecting, adding to their own knowledge and translating their thoughts into language.

The information is gained by the recipient by direct transmission or a retrieval process and received via verbal or visual stimuli. They respond to the information in one of the ways discussed above and potentially add it into their own knowledge. In order to get at the information, a recipient must consult their own knowledge, find a gap in it and translate that gap into a question that can be used to gain the information.

This chain, the relationship between information and knowledge and originator and recipient, Farradane saw as being crucial to LIS. An originator turns knowledge to information and a recipient back again (Farradane 1980) but language acts as a barrier in doing so and, for this reason, Farradane felt these relationships were important and that users therefore needed more attention as they sought to frame questions in a way to get at the information required.

2.4.3.3 Nicholas Belkin (and Thellefsen, Sørensen and Thellefsen)

Nicholas Belkin took a similar view to Farradane (1979 and 1980) in some respects. His ideas were first put forward with Stephen Robertson in Belkin and Robertson (1976). Here they define Information in Information Science as “the structure of any text which is capable

of changing the image-structure of a recipient." This "image-structure" can be taken to be the knowledge of an individual which, as with Farradane (1979 and 1980) and Brookes (1980), is altered by information.

Belkin (1978), though, is more specific in his definition of information which he says "must be purposeful, meaningful… It must have an effect on the recipient. It must affect the knowledge state of the recipient" (Thellefsen et al 2014). As Thellefsen et al (2014) point out, information can only be information if "the generator is capable of communicating information in a purposeful and meaningful way." (As Shannon and Weaver might see it, perhaps this equates to a lack of entropy and creating the correct message. Or, as Farradane might see this, language can act as a barrier to something being meaningful, either because you don't understand the language or because of poor construction).

Belkin and Robertson (1976) additionally list three phenomena of Information Science:

I The texts and its structure (the information)

II The image-structure of the recipient and the changes in that structure

III The image-structure of the sender and the structuring of the text

Here we also see the term text, which could be seen as synonymous with data, if only because text is presented as separate from information (which is defined as the structure of texts, texts being the symbols that make up that text). However, they do not expand on the relationships other than stating their existence.

In Belkin (1978) these ideas are developed further, however. While discussing the information concept and, via a discussion of previous work, Belkin comes to the conclusion that the main problem of information science is "facilitating the effective communication of desired information between human generator and human user" and that information scientists should therefore have the following concerns:

1. with information in human, cognitive communication systems;
2. with the relationship between information and generator;
3. with the relationship between information and user;
4. with the idea of desired information;
5. with the effectiveness of information and the effectiveness of information transfer.

Furthermore, Belkin also creates a list of features that should be included in an information concept for information science (he believes this concept is more important than defining information *per se*), which are shown in Figure 9, below:

1. It must refer to information within the context of purposeful, meaningful communication. (D)
2. It should account for information as a social communication process among human beings. (D)
3. It should account for information's being requested or desired. (D)
4. It should account for the effect of information on the recipient. (D/B)
5. It must account for the relationship between information and state of knowledge (of generator and of recipient). (D/B)
6. It should account for the varying effects of messages presented in different ways. (B)
7. It must be generalizable beyond the individual case. (M)
8. It should offer a means for prediction of the effect of information. (M)

D Definitional Requirement
B Behavioural Requirement
M Methodological Requirement

Requirements 1–6 are relevance requirements
Requirements 7 and 8 are operational requirements

Figure 9: Belkin's Features for an Information Concept (Belkin 1978)

Between these we can see the similarities between Belkin and Farradane's ideas in that each are concerned in the relationships between generator and recipient and information and knowledge. Indeed, Belkin (1978), while defining science talks about it "attaining a certain kind of knowledge," suggesting Belkin and Farradane are also both concerned with the discovery of information and its processes.

Belkin's (1978) ideas have been discussed further by Thellefsen et al (2014) who say Belkin addresses the effect of information on a recipient and how it becomes knowledge- issues they think were not developed further after Belkin's article.

From a semiotic viewpoint and using their own significance-effects (emotion, information, knowledge), Thellefsen et al look further at how information can become knowledge through the example of a walker seeing a dog swimming in the lake from a distance.

As the walker gets closer and gains a better view, each significance-effect becomes more dominant until they are able to fully ascertain the situation. What starts as a dot on the horizon, allowing emotion and imagination to run riot, later becomes an unspecified dog, as more information (through the number of characteristics seen) is gleaned; until the walker is finally able to see, using their knowledge, that the dog is, in fact, their neighbour's dog, Jason (see Figure 10 for the article's summary).

Semeiotic	Levels of meaning creation	Hypotheses	Significance-effects
A spot in the horizon	Emotion (dominant) Information Knowledge	The number of characteristics is low, and the number of hypotheses is high	Emotional significance-effect
An unspecified dog	Emotion Information (dominant) Knowledge	The number of characteristics is increasing as the number of hypotheses is decreasing	Informational significance-effect
Jason, the neighbor's dog	Emotion Information Knowledge (dominant)	The number of characteristics is high and the number of hypotheses is low	Cognitional significance-effect

Figure 10: Thellefsen, Sørensen and Thellefsen's Elements in the meaning creation process (Thellefsen et al 2014)

From this example, Thellefsen et al (2014) demonstrate how data, information and knowledge interact together, each needing to be reassessed and interrogated as more is known.

At the start, the walker knows very little about the dog and is able to imagine it as a bear. At this point in particular, different people might come to a different conclusion from what they can see and have previously experienced. As they get closer and gain more information and use their own knowledge more, they can see that it is a dog and then find they know the dog personally.

Thellefsen et al (2014) say cognition brings together emotion and information, blending the external and internal worlds (see Figure 11, below). Cognition mediates between emotion and information to ultimately bring someone to the correct result.

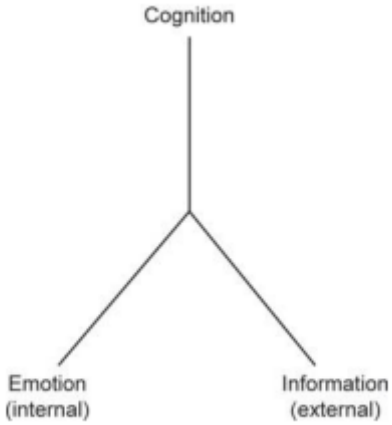

Figure 11: Cognition bridging the gap between emotion and information (Thellefsen et al 2014)

Between them Belkin and Thellefsen et al present the relationship between information and knowledge as one of cognition, which the latter show to be complicated by emotion.

2.4.3.4 Data-Information-Knowledge-Wisdom's Progression and Ma's Intricate Web

An important and often discussed model of the relationship between data, information and knowledge is the Data-Information-Knowledge-Wisdom model (DIKW). The model was probably introduced to LIS in 1988 via a paper given by Russell Ackoff, though Ma (2012) says it dates back to at least the 1960s. Ackoff (1989), though, did not lay it out diagrammatically but people have subsequently expressed the model in different ways, two of which are:

in a linear form:

Data - Information - Knowledge - Wisdom

or a pyramid, reflecting the hierarchical nature of the model (Rowley 2007):

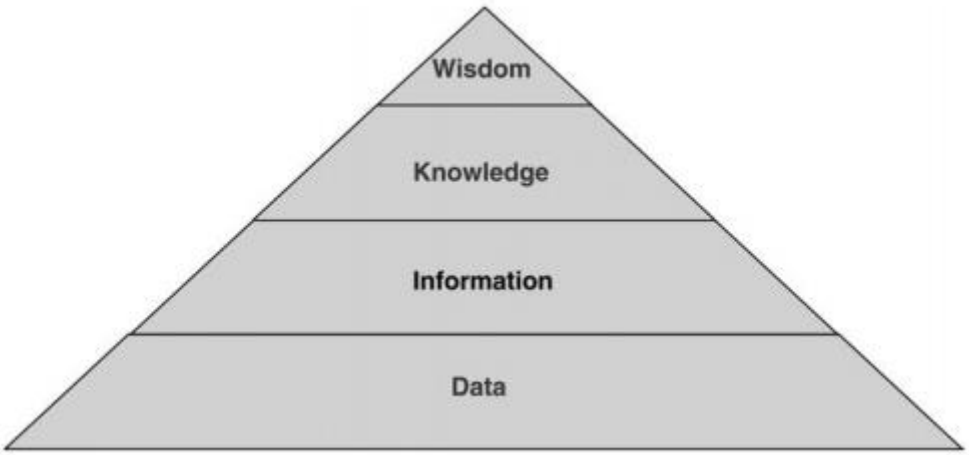

Figure 12: The DIKW hierarchy shown as a pyramid (Rowley 2007)

DIKW has received much criticism since its inception. Ma (2012) feels it is limited because it deals so much with computing and gives a systems view of the world. (Ackoff (1989), indeed, refers constantly to Management Information Systems and improving workplace efficiency).

The model, Ma (2012) says, looks at the progression from data to wisdom in the human mind as a series of data-processing events that represent human thought and learning. This, she says, is too limiting, and by quoting Machlup (1983), points out that a sender and receiver are needed to impart information.

Ma goes on to show that information and knowledge are not processed data: "rather, we learn by being situated within and understanding complex webs of relations of persons, events, social and political structures, and many other things." In this web, adds Ma, we seek agreement from one another "on what things mean through learned social and cultural tools and categories," and points out that human learning is developed over time whereas computers process data according to a fixed programme.

She concludes this criticism with a rather lovely sentence: "The analogy of data to human stimuli and machine process output to knowledge is a bizarre analogy that obfuscates rather obvious difference between designed and organic agents."

As with her criticism of Brookes' equation and use of Popper's Worlds, Ma sees DIKW as being too simplistic and shows that the relationship between data, information and knowledge is much more complex than a simple process from A to D (or D to W).

Much of what Ma (2012) says is backed up by Martin Frické (2009) who believes that the DIKW pyramid is built without foundations. This conclusion comes from Ackoff's definition of data as "low-level true statements." Frické says this makes DIKW "conservative" in regards to the nature of information, which means that it cannot allow for inferred information or statistical generalisations which may not necessarily be true. DIKW also does not allow one to ask "why," it forbids you to go beyond the data and find out what is happening; and so it does not allow for information that has not been derived from data, i.e. internally from old knowledge and inferences.

Between them Ma and Frické raise some important points that help point toward the relationship between data, information and knowledge that are different from the simple linear progression of DIKW.

2.4.3.5 Marcia Bates and Birgir Hjørland's Interpretations

Across two articles (Bates 2005 and 2006) Marcia Bates discussed her views on the definitions of data, information and knowledge. In doing so, she also talks about the relationships between them.

In Bates (2005) data can be two things: Data 1 is a "portion of the entire information environment available to a sensing organism that is taken in, or processed, by that organism." Anything outside their sphere is Information 1 and the Data 1 that animals perceive becomes Information 2 (see below for Information 1 and 2 definitions). Data 2 is "information selected or generated by human beings for social purposes." This second data ultimately helps account for the higher thinking of humans compared to animals. Where Data 1 relates more to environmental factors and survival, Data 2 can be as simple as a child touching a grasshopper to see what happens to data gathered for research.

Bates extends the meaning of information presented by many LIS writers by giving us Information 1 (the pattern of organisation of matter and energy) and Information 2 (the pattern of organisation of matter and energy given meaning by a human being) (Bates 2006). By doing so, Bates says that information is in everything whether it is interacted with or not and therefore that it has no inherent meaning.

In saying this, she is in opposition to many writers, in particular Bateson, whom she discusses (Bates 2006). Bateson said information is "a difference that makes a difference." "To whom or what?" asks Bates, saying we should begin before that understanding and, thus, the suggestions of Information 1 and Information 2. The pattern of organisation within all things is information, as is our later interpretation of it. Any physical material has the potential of carrying any meaning and to inform somebody (Bates 2008).

Knowledge, Bates (2005 and 2006) defines as "Information given meaning and integrated with other contents of understanding." This is entirely brain-based and "dies with us," as knowledge contained within books is Information 1 (Bates 2005).

In this, Bates gives us the relationship between data, information and knowledge. Or, perhaps, gives us the relationships for humans as well as for animals. Ultimately, what Bates presents is a process that takes sensing organisms from collecting the Information 1 that surrounds them in the form of Data 1 or Data 2, processing them to form Information 2 and using this to add to their own Knowledge store. Whilst possibly being more complicated (maybe over complicated; I haven't even touched on the vast range of Information types presented in Bates 2006) and wide ranging than some writers, what Bates posits isn't far removed from other models presented here thus far.

Birgir Hjørland, in his opposition to Bates (Hjørland 2007) is clearly in the Bateson camp. He does not like that Bates thinks that Information can be any difference, even if it makes no difference, and says that such a definition makes information too vague and useless.

Hjørland sees information in signs and signals which can mean different things depending on an individual's knowledge base and therein he sees the relationship between the two. This is something that Hjørland had already written about a lot before his debate with Bates through his work on social epistemology, which Hjørland said information is a type of (Bonnevie 2001).

Hjørland (1998) makes an argument for IS to bring epistemology into the fold, saying it has a close relationship with the subject and can be learned from. Information theory, including Shannon's, though, are a part of computer science and not IS.

Across several articles, Hjørland (2002, 2003, and 2007) has continued this argument, stating the importance of knowledge in information science. Within these he talks about the need to have knowledge in order to make something informative (Hjørland 2007). He also writes about knowledge within systems such as information retrieval systems, without which we would not be able to get at information. We also need knowledge in order to use these systems: an "elementary, yet rather neglected fact" (Hjørland 2002). These aren't the same types of knowledge, as Hjørland states, but are each of vital importance within information science.

Birgir Hjørland therefore sees knowledge as more important than information, and as something that is not just the end result of a process or something that is altered by information, but something that is far more involved and that has an effect on knowledge.

2.4.3.6 The Views of Madden, Buckland and Machlup

Madden (2000) talks about the relationship between data, information and knowledge as one that rests upon the contexts in which messages are received and sent. Madden merges data, information and knowledge together in some ways, saying that information can be passed on in various forms:

- as a representation of knowledge (or stored knowledge in books etc.);
- as data in the environment (information gained from nature, for example, such as the markings on a butterfly's wings);
- as a part of the communication process ("Meanings are in people rather than in words or data," here Madden talks of how timing and social factors affect information); and
- as a resource (information is transmitted from a sender to a receiver and may have additional meaning as it is disseminated).

To this Madden adds that messages containing information are affected by the readership and authorial contexts, i.e. the receiver and sender of the message. The information received can be interpreted through various factors or filters such as Geographical, Social, Educational and Professional. To these factors, that of intention, can affect the message from the sender's point of view when composing it.

Added to this, Madden brings one's own knowledge into play; firstly as it is needed to recognise the context of a message and filter information for meaning. But also, in relation to information seeking, Madden points out that users will search for information based on a gap in their current knowledge set. And this knowledge set will cause a recipient to determine whether information is sought, what information is sought, and how it is interpreted. This knowledge will additionally be made up by the four filters mentioned above.

Buckland (1991) sees information in various forms, two of which include data and knowledge. His main forms, though, (all of which hail from the Oxford English Dictionary and identifying common usages of the word) are:

information as process - the act of being informed
information as knowledge - that which has been perceived in the process (this can increase uncertainty and is intangible
information as thing - objects, including data and documents, that impart information

As was shown in Figure 2, Buckland (1991) divides information into tangible and intangible entities and processes which interact to form a relationship. Buckland (1991) does not explicitly talk much about the relationship between data, information and knowledge (indeed, he is mostly only concerned with definitions of information here) but this figure does give some insight into a four way process linking tangible, physical things, such as documents

and information processing with the intangibles of knowledge and communication. Buckland, here, presents a chain of communication or events that bring together data, documents and knowledge under one banner of information types. This could be seen as a less structured version of Todd's interpretation of Brookes (Figure 7).

Buckland (2012b) approaches this chain a bit more by seeking to focus information in IS to human knowing and learning, thus focusing the information-knowledge relationship on communication.

In his concluding remarks, Buckland states that Information Science should be about enabling people to learn, to become better informed and more knowledgeable ("information services are, in practice, more directly concerned with *knowing about* than with *knowing how* or *knowing that*." (emphasis in original)) and, in doing so, talks about the importance of document selection, especially in a world in which people have no time to seek raw information and have to rely on secondhand information in the form of documents.

Thus, for Buckland, it would seem that the relationship between information and knowledge that should be important is the knowledge to find and select the information required.

In an often cited article from 1983, Fritz Machlup discussed the *Semantic quirks in studies of information.*

Machlup believed that, for information to be imparted, two people were needed: a sender and a receiver (Ma 2012). Information, he said, is acquired by being told (Machlup 1983) while knowledge is acquired by thinking. This has the interesting knock-on effect that "new knowledge can be acquired without new information being received," which is something that is normally ignored in most information-knowledge relationships.

Information is a process and knowledge is a state, adds Machlup (1983), adding that modern parlance helps us see the difference: we would expect to find an information stand, for example, not a knowledge stand at somewhere such as a railway station, as it is a place where one undertakes the process of being informed.

Machlup also uses the metaphor of observation in science to demonstrate the information process. Whether it is an experiment or an interaction with two people, one who tells and one who listens is required (instrument readings can stand in for one of the people). One or the other on its own is just one part of the process- it is only when something is relayed on that we get information.

2.4.3.7 Dervin's Three Informations

Brenda Dervin (1977) puts forward an interesting threefold definition of information, which include data, and arguably contains knowledge as well. Her third definition brings in an extra element that is not often integrated into data-information-knowledge relationships.

$Information_1$ and $Information_2$ in Dervin (1977) are defined as external or objective and internal or subjective information respectively, with $Information_1$ being defined as "information which describes reality, the innate structure or pattern of reality, data" and $Information_2$ as "ideas, the structures or pictures imputed to reality by people." Arguably, both these definitions could be defining information as synonymous with knowledge.

$Information_3$ is finally defined as the behaviours that link $Information_1$ (the information an individual selects) and $Information_2$ (the information an individual creates). These behaviours, ultimately a decision making process, can take several forms, says Dervin, including an individual taking stock of a situation or asking others for help. $Information_3$

could also mean the knowledge processing that occurs when a student writes an assignment.

Although, not all the three terms are used, Dervin is describing a process along the lines of a relationship with data, information and knowledge. In this relationship, the external and objective $Information_1$ required are selected by an individual before being converted by the behavioural $Information_3$ into $Information_2$.

3. Discussions

3.1 Introduction

In the previous sections, a view of the definitions of and relationships between data, information and knowledge from LIS have been presented. From these, it can be seen that there are various interpretations of these concepts and the way they relate to one another. The following sections highlight the main issues arising from this review.

3.2.1 Data-Information-Knowledge is a process, but how does it work?

The relationship between data, information and knowledge is normally presented as a process. The exact meanings of each part of that process and how each becomes the next within that process may differ somewhat but all writers agree that we are dealing with a change.

One of the key differences here is the complexity of the process. For Shannon-Weaver, Brookes and DIKW, the process is quite simple (Ma 2012): something goes in, something comes out. In Shannon-Weaver, a message enters a system and we hope to get the right message out; in Brookes (1980), information is added to a pre-existing knowledge state to create a new one; and in Ackoff (1989), an individual processes data up the chain to eventually become knowledge.

Other writers (Machlup 1983, Farradane 1979 and 1980) see this process more like a two-way conversation between a sender and a receiver- either directly or via a text or document. In these cases, information is transferred by discourse and added to the receiver's knowledge. Such an idea seems to cut out data and only refer to the information-knowledge relationship. However, I do not think this is always the case, as Machlup 1983's definition of data as what is given allows for the work of a researcher gathering data or to see the initial transfer of a message (whether spoken or in writing) as data before being translated into information through translation of its language.

Although he starts by saying some studies give it undue prominence, Farradane (1980) talks about language as the medium through which humans have to transfer thoughts to one another in as we "cannot communicate by direct thought transference." This language can be in nature "mathematical, graphical or pictorial, or some art form such as paintings or music."

Written and spoken language cannot fully translate thoughts and the rules that govern it do not let it flow in the same way as thoughts do, but humans have an innate skill at learning language (which has to be learned, we are not born with it) and we use it to rationalise our sentient experiences (Farradane 1980).

This is ultimately automatic and is a process we perform without thinking about it very much at all. But it is an important part of the data, information, knowledge relationship that is perhaps often overlooked. It is something that McCloud (1993) explains neatly too in showing how images within comics create their own language that needs to be learned for one to fully understand.

A further example is presented by James Gleick (2011) who begins his look at Information through the ages with the example of the African drumming language recorded by John F. Carrington in the 1940s. The language is communicated by tonal drumming that matches the tonal language in Zaire, where Carrington lived. This had made the language very difficult for Europeans to understand. Once Carrington had worked it out, though, it became

simple: because of the tonal nature of the language, more information would be required in the message to add the context that was the key to its understanding.

Farradane (1980)'s model (see Figure 8, above) brings language issues into the data-information-knowledge relationship and expands somewhat on the simpler idea often purported of information passing between a sender and receiver by mapping out the whole process.

As reported above, Ma (2012) in criticising the simplicity of Brookes, Shannon-Weaver and DIKW, sees the data-information-knowledge relationship as an intricate web. Ma is just one of the LIS writers who sees the relationship between data, information and knowledge as being more than a simple one or two-way process.

Wiener proposed a feedback system in information, demonstrated by Claude Shannon's rat learning to get out of a maze (Gleick 2011) while Brookes (1974) talked about gradually building up knowledge while taking in information bit by bit. While Thellefsen et al (2014) describe a back and forth interrogation of information in order to gain more data and decrease uncertainty.

They are not alone in such an idea. Ford (2004) looks at the problem-solving feedback model of Wilson in which Wilson showed how someone will resolve a problem through a four stage process of: problem identification, problem definition, problem resolution, and solution statement. "Where the attempt to resolve uncertainty is not successful at any stage, a previous stage may be revisited" (Ford 2004). In other words, a gap in knowledge is identified, a way to solve it is theorised, information is sought and an answer found. Wilson said that, at any stage, someone might double back and repeat a step if the uncertainty from their original knowledge gap was not filled.

3.2.2 The Importance (or Otherwise) of Truth

Whether or not data, information and knowledge need to be true is something that divides the field a little bit, with most writers seeming to suggest it does not need to be. As mentioned above, Floridi (2010), sees semantic information as being well formed, meaningful and truthful data. Knowledge, then, is relevant semantic information properly accounted for.

DIKW also sees truth as being fundamental to its definition of data, which Frické (2009) believes weakens the idea. Many LIS writers agree that truth is not required, seeing knowledge as a kind of meaningful belief (Bates 2005). Zins (2006) phrases this differently, seeing (universal) knowledge as the meaning or content of thoughts the individual justifiably believes to be true. As Buckland (2012b) says, no one would say their beliefs are unjustified, meaning that the justified part of the philosophical approach to knowledge does not work. Buckland also says that truth cannot work as knowledge is always based on some other prior belief or assumption.

This keys in with the premise of the television programme *QI*, which is based around debunking popular beliefs or knowledge and shows that knowledge is fallible and will often be untrue (or, to be kind, just not up to date). Machlup (1983) adds that you should always question the truth in something, quoting Bertrand Russell who said "all human knowledge is uncertain, inexact and partial."

As shown by Frické's (2009) critique of DIKW, whether or not data, information or knowledge need to be true has an effect on the relationship between the concepts. On a personal level, I would agree with those LIS writers (and Bertrand Russell) who see knowledge, in particular, as not needing to be true. I often discover that my knowledge is wrong or find

information that is wrong. An important part of the relationship between data, information and knowledge is to continually question the latter by sifting new data and information, or thinking and hypothesising.

Floridi (2010) separates out misinformation and disinformation from true factual information, saying that they cannot become knowledge. As stated above, I do not agree, as no person's knowledge base, and no physical store of knowledge, will ever be completely correct. What is in your mind could be anything, but what Floridi is stating is what we can strive for, and what, indeed, subject, or society, knowledge will strive for. The work of academic disciplines is to try and get close to the ideal Floridi has set.

Perhaps, in this area, it is better to use Belkin (1978)'s idea that information must be purposeful and meaningful, rather than truthful, to become knowledge. Certainly, it is more of a barrier for information to be poorly structured in its language for it to be taken onboard, than it is to be false. Ultimately humans will remember, or believe, any old rubbish so long as it is easy to understand.

3.2.3 Where should knowledge be?

Knowledge and its place, or places, in its relationship with data and information is another issue that this literature review has highlighted. It could be seen as the obvious thing to always place knowledge at the end or top of a process chart: the end result from data and information, or the state that is altered by them, as in DIKW or Brookes' fundamental equation.

But this is not always the case. Todd's (1999) rendering of the ideas of Bertram Brookes places knowledge (in different forms) at either end of a process, and Farradane (1980) places knowledge in several places (and again in different forms) in the back and forth between a sender and recipient of information (including the two points of language translation).

By seeing the relationship between data, information and knowledge as an intricate web of learning, Lai Ma (2012) suggests that knowledge in different forms and from different sources (including people, events, society and political structures) may be continually interacted with as one learns.

A similar view could be seen in Shera's social epistemology or Hjørland 's domain analysis (Bawden and Robinson 2012) in which knowledge is shared across society or a particular subject area and drawn upon to help gain or understand information. While Thellefsen et al's (2014) cognitive example shows knowledge being constantly interrogated by information to reduce uncertainty.

As discussed above, Farradane (1980), McCloud (1993) and Gleick (2011) each talk about the need for language to understand data given before it can be turned to information. This again shows a further point where knowledge may be placed first rather than being an end product, as understanding, or knowledge, is required to process data or information. As Hjørland put it, understanding is determined by our pre-understanding (Hjørland 2002) and knowledge is needed to make something informative (Hjørland 2007). In other words, we need knowledge (or knowledge of a gap in it (Belkin et al 1982)), before we can start the process of gaining new knowledge.

Similarly, Zins (2004) talks about how already existing concepts can cloud what we think about a piece of information or data. The examples he uses are pictures of the New York skyline and the Star of David, saying that the knowledge he already holds will control his thoughts about them. For him, the New York skyline makes him think about terrorism and

the Star of David Judaism, neither of which are explicitly shown in the photos. Other people may think of other things. Again, however, we see how knowledge could be seen to come before information or data. Rather than pre-understanding, this is pre-judgement or bias, a topic touched on above by Farradane (1979).

In Atkinson (2014) it was suggested that the notion of dust as presented in Philip Pullman's *His Dark Materials* novels (Pullman 1995/2011, 1997/2011 and 2000/2011) might provide a good analogy of information. It may be that, as in the books themselves, it may represent knowledge better in this overarching aspect as something that is always there to be interacted with. This view will be discussed further in Section 4.1.

3.2.4 Computers vs Humans

Some divisions in ideas about the relationship between data, information and knowledge comes down to whether or not computers are a part of the process. As noted above, Machlup (1983) talks about data both as anything given to a researcher but also anything fed into a computer.

Elsewhere, Shannon's Mathematical Theory of Communication is ultimately concerned about engineering and DIKW was initially conceived as a way of looking at Mature Information Systems; arguably lending each idea better to computer studies.

Brookes, however, among many others, is concerned only with the interaction between humans and what happens in their minds. Bates goes further and extends information processing to creatures other than humans.

This could be argued to be partially the difference between the L and I of LIS, or, at least because it is a combination of computing and social sciences. As both a student undertaking a Masters degree in Library Science and an employee in a university library, it has often struck me that there are these two sides of LIS: the more physical libraries side where we deal with the humans and the resources seeking to bring information to users and the computing side that underpins all of this (the systems behind the Library Management Systems and OPACs, for example).

In terms of staffing, libraries employ systems and metadata librarians who deal with the inner workings of the systems used by staff and patrons, while customer services and subject librarians help patrons use these.

During the course, such a division was perhaps most strikingly shown when Library Science students took a module called Digital Libraries whilst the Information Science students took Information Retrieval. Each looked at library systems but while the Library Science students looked mostly at them from the user's point of view, the Information Science students delved much further beneath the hood.

Similarly, an exhibition at the Science Museum in London, called The Information Age, looks at information in the same way as James Gleick does in his 2011 book, *The Information*: i.e., more through the development of science and the artefacts that helped convey information (or "more than 200 years of innovation in information and communication technologies," as the Science Museum website says). Rather than looking at the story of documents and libraries, each focus more on things like calculation machines, the telegraph and computers as conveyors and processors of information (although Gleick does also look at early documents). Arguably the British Library might view The Information Age from a different angle entirely.

Each of these viewpoints may look at the data, information, knowledge relationship differently. Especially as computers have become more and more involved in libraries and the importance of digital skills has increased. Should we be worried more about how the human computer interaction is involved when humans process data, information and knowledge?

Bawden (2008) reports that Meadow saw information science as becoming an "integrating science" in the 21st century. Perhaps these two different views of the relationship between data, information and knowledge should be a part of the integration.

3.2.5 Work from other disciplines, or Is everyone only in it for themselves?

Often, but by no means always, LIS writers look to other disciplines to help resolve their issues with the definitions of and relationships between data, information and knowledge. In this particular part of the discipline this makes a lot of sense, especially in regard to the definition of information which can have quite different meanings in different disciplines (Capurro and Hjørland 2003).

Examples from what has been discussed here include Bates (2006, Atkinson 2014) drawing on the work of evolutionary biology and Hjørland and Zins drawing on philosophy, in particular epistemology. Floridi (2010), a philosopher of information, looks at his subject from the viewpoint of information in various disciplines. In his book he covers Shannon-Weaver/MTC from the field of engineering, before considering mathematical, semantic, physical, biological and economical information.

Robinson and Bawden (2013) followed this up with a comprehensive review of the concept of information in different disciplines, concluding that there are gaps between these concepts of information in different disciplines and that LIS should try to fill these: "Mind the gaps, certainly, but be aware of the insights that may be found within them."

I do not seek to propose that LIS writers should not look away from their discipline to seek answers for problems, or at least help in doing so. It will always be infinitely interesting to compare and contrast views to fuel discussions.

Potentially, though, a closer look at the current LIS discipline might be of use too. Farradane, Belkin and Brookes all approached information while thinking about information retrieval systems; I would like to do something similar here with Information Retrieval, Information Literacy and Knowledge Management.

First, however, I would like to discuss further issues with looking at other disciplines to solve problems. To conclude this section, I would like to reflect on a point quite often made about articles such as those that have been reviewed above, including by myself (Atkinson 2014), that literally any viewpoint can be argued if you can find the right article to back you up.

Something that could be becoming more and more possible, especially in universities, as discovery systems become more commonplace in libraries (alongside database searching platforms such as Ebscohost) and Google Scholar becomes more widely used.

Information theory has been criticised several times along these lines.

Machlup (1983) summed this up in regards to definitions of data, information and knowledge by saying, "one can probably find quotations supporting all possible combinations of the three terms or of concepts they are supposed to denote."

Atkinson (2014) similarly criticises the Bates-Hjørland debate (and academia in general) by pointing out that each make arguments by finding a view that fits. Where Bates (2006) uses an evolutionary biologist to support her view of information, Hjørland (2007) finds one to support his, while "Buckland and Hjørland both use computers to create different views on information" (Atkinson 2014). Bawden (2001) likens this to Wittgenstein's "language games" in which the meaning of a term can be assigned by its context

Hjørland himself wondered about this idea when he pondered whether theories from other fields and 'unconscious attitudes' have guided research and practice. With such unconscious attitudes, could LIS writers be ultimately writing the theories that suit them?

As such, I think it would be useful to look at examples within LIS to see if they help shed light on the relationship between data, information and knowledge.

3.3 Data-Information-Knowledge in Individual Fields of LIS

3.3.1 Data, Information and Knowledge in Information Retrieval (IR)

Information retrieval systems have been in existence from the 1960s and a system called Dialog (Bawden and Robinson 2012, MacFarlane 2012/2013). They are used as an aid for people to find the documents or information they require and can vary massively in how complicated they are to use from systems with advanced searches using Boolean terms to Google, which has one search box and can find documents/web pages with one word (Bawden and Robinson 2012).

Consulting an information retrieval system is a little like having a short conversation with a computer in that a user will submit a query to be answered. In this case the answer will be a list of documents or records for documents that the system deems to be relevant. As this is an information exchange, various LIS writers have written about theories of information in relation to IR (Farradane 1980, Belkin 1978).

A central concept of an IR is that it will be used by someone seeking to fill an anomalous state of knowledge (ASK; Belkin et al 1982). This ASK is then used to form a query that a user will enter as a search in order to satisfy that ASK.

That search can be entered in a variety of ways, often using Boolean terms such as AND, OR and NOT in order to phrase a query appropriately. This can be quite strict or using several search boxes linked by Boolean phrases (see examples in Chapter 7, Bawden and Robinson 2012). More natural language systems such as Google will employ algorithms that take a search query and use it to generate results based on relevance: for Google this is to do with the number of websites that link to the pages displayed in the results (Bawden and Robinson 2012).

The results will often be displayed in a list of documents that can either be viewed through the same system (the search tool itself or the browser when web based), if the items are digital, or need to be searched for if the items are physical, the information returned giving its location.

Throughout this process, knowledge of various kinds, beyond awareness of the initial ASK, in order to gain the data is required. Farradane (1980) talks both of the structure of the language required to create a search as well as that required for the indexing of documents. Each of these are languages, or knowledge bases, required both from the user and system points of view (MacFarlane 2012/2013).

Gärdenfors (quoted in Hjørland 2002) talks about there being knowledge in the system while Hjørland (2002) says an IR user needs to have knowledge of the four levels they are interacting with, which are:

(1) … a personal computer (PC or MAC) and its operating system.
(2) … a remote computer system (e.g., Dialog) and with its specific search engine and file organization.
(3) … with document representations (e.g., bibliographical records).
(4) … representations of the subject literature in one or more domains (which can be more or less homogeneously or merged).

Although in this quote Hjørland talks of “document representations” and had earlier stated that IR deals with documents not information (Hjørland 1999), this at the very least can be seen as data, if not the information required to get at the information sought.

Hjørland (2002) also notes that IR is not just about the expressiveness of natural language in a query but a matter of real knowledge of what is being searched for - but also that this knowledge is not the same as subject knowledge.

Ford (2004) talks about theory generation and the use of tacit unconscious knowledge to obtain the information required, presenting various figures showing the processes, including one reproduced here as Figure 13, which shows the internal dialogue of an information seeker.

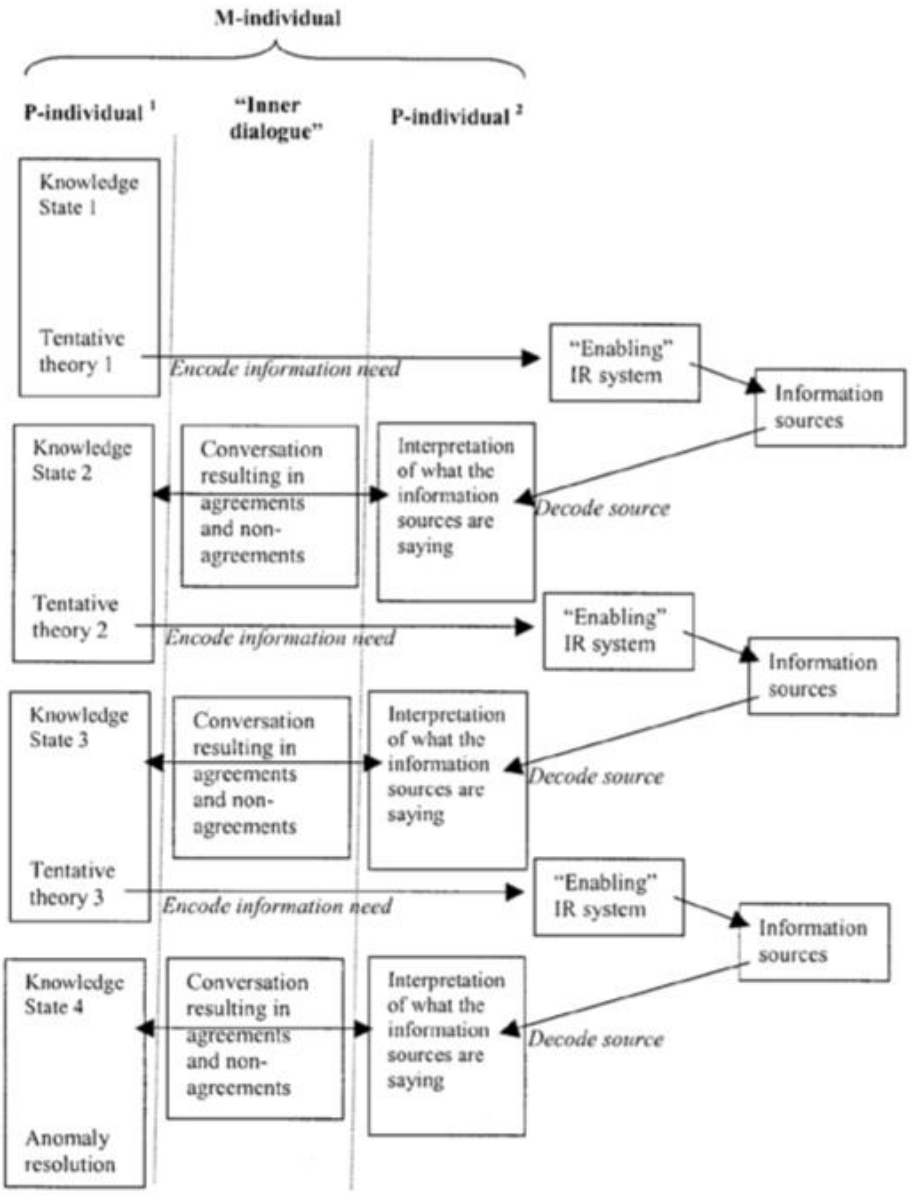

Figure 13: Ford (2004)'s model of the internal dialogue of an information seeker.

This is similar to Wilson's feedback loop and shows how people will try different questions and searches in order to find what they need, continually consulting their tacit knowledge to get information.

This is all demonstrated by my own actions while conducting research for this dissertation. I used the Ebscohost platform to search a database called LISTA for relevant articles. To do so I worked out a search using terms from my question, set up a series of searches for different sets of terms using the Boolean operator OR; and then used the Search History to combine those searches using the Boolean operator AND.

I was given a list of documents that, from the way they had been indexed, matched my search terms. By looking at the abstracts of these documents, I was able to determine what would be useful/relevant to read and use in my work.

I also used CitySearch, City University London's discovery system (Serial Solution's Summon) which searches their own Library Management System (Innovative's Sierra) as well as all the electronic resources to which the University library subscribes or owns. By using a more general, or subject search, I was given a list of articles and books that matched my query.

In both cases, I had an ASK (I needed articles and books that would give me information to make me more knowledgeable about the relationship between data, information and knowledge) but alongside that gap in my knowledge, I also had the knowledge required to fill it. I knew where to look, and how to look. Once I gained my results, I had the knowledge to refine the results and assess what I was given before reading the articles themselves.

It was only after applying that knowledge that I gained the information (or the means to access it) that would ultimately fulfil my ASK and add to my knowledge state. Additionally, it was only due to the system being set up with knowledge by individuals that I was able to get the correct results (the knowledge in the system).

Without people cataloguing the books and articles, adding subject headings, writing the abstracts, i.e., using intellectual tools and systems such as Library of Congress Subject Headings and MARC to add to the documents, I would not get any results at all.

This was something that the Soviet Information Scientist, IU A.Shreider, saw as being very important indeed. Schreider saw "information… as an estimate and/or characterization of knowledge, either content knowledge or missing (that is, needed) knowledge" and the problem of informatics as being what someone can learn from a text (Belkin 1975).

Consequently, Schreider saw the relationship between information and knowledge that information scientists should be concerned with as the metainformation, or knowledge of a text, that is used to find texts. This, then, is not a new idea, but is an important thing to pick up on here.

Overall, then, I put data/information into the system and get data/information out of the system, which I can use to get the documents required, which will give me information to add to my personal knowledge store.

However, overarching all of this, I am constantly using my own knowledge, understanding and experience. Firstly, to decide on the information that I need and secondly to use the system and assess the results. If I found the results to be unsatisfactory, I would try different searches until I was successful.

I followed both Hjørland and Ford's ideas not just by using knowledge I had and knowledge in the system to get at the information, but by trying different ways to find the answer.

Discovery tools CitySearch and Google, I think, encourage the learning of more flexible or direct ways of searching; or the use of experience. For example, while writing this dissertation, I would often return to articles via CityLibrary Search to look at them again. After doing so a few times I learned the best ways to ensure I quickly found the desired article. For example, a search for "shifting Bawden" was enough to get me Bawden (2001). In other words, people quickly build this knowledge for finding the information they need.

There is no simple move from data to knowledge, therefore, in IR. The process starts and finishes with knowledge (once the documents retrieved have been read), with the knowledge of and in the system coming in between.

3.3.2 Data, Information and Knowledge in Information Literacy (IL)

Skills such as those used in IR are learned through information literacy (IL). Indeed, as Bawden and Robinson (2012) show, the skills set up above are what make a person information literate, adding to them the effective use of the information gained (they quote the American Library Association from 1989 who said: "Information Literacy is defined as the ability to know when there is a need for information, to be able to identify, locate, evaluate, and effectively use that information for the issue or problem at hand.")

There is another side to IL, however, and that is the teaching of it- the enabling of people to use the resources available in libraries. There is the building of the skills and the understanding in the first place.

This is demonstrated in schools and universities where information literacy is taught in the libraries. Many have information literacy strategies in place to help with this, providing 1-2-1 and group sessions which are presented in various ways.

The Information Literacy Strategy at the University of Salford (University of Salford 2011), for example, gives 8 Levels through which students can progress and learn the skills required to complete their courses (and which they will use for the rest of their lives). Level 1 contains the skills that new students are expected to have, but may need to be taught on arrival, while Level 8 are the skills required of a PhD candidate.

Such a build-up in information skills learning supports many of the ideas discussed above, including the building up of knowledge through the gaining of information skills (Brookes 1974).

The transfer of skills will occur in group or individual training. My own experience in this area mostly relates to 1-2-1 sessions with health science students at City University London. In these sessions, I would talk a student through how to perform a literature search, usually for a piece of coursework they were just starting out on.

In terms of a data-information-knowledge transfer, such a session would follow more of a linear view as I sought to transfer my knowledge to the student for them to use in future. In doing so I would turn knowledge into information through communication. I would be the sender to the student's recipient.

Where knowledge might play a slightly different role here is in regard to understanding and might be better demonstrated in regard to the Shannon-Weaver model of information transfer.

No teacher is successful if their students do not understand what they are being taught and will leave without any new knowledge being gained. In Shannon-Weaver's model of communication, noise is shown to be the main factor in potentially meaning a wrong message is received.

For Shannon, this noise was interference on a telephone line. In face to face communication, noise from those around could potentially play a factor in a message not being received correctly. For example, in a noisy pub or club where a bartender is unable to hear an order properly because of multiple other conversations occurring at the same time in close proximity (when I was a teenager, the main danger in such a situation was when ordering a bottle of Beck's or Metz: each one being easy to mishear as the other).

In a literature skills session 1-2-1, the noise comes from a mixture of a student's understanding (related to previous experience, i.e. where they are on a scale such as in

Salford's IL strategy) and the way the teacher is teaching. Where these do not meet up, noise occurs and the message will not be fully received and understood. I would seek to combat this by trying to gauge the information literacy skills of the student at the start of the session and then tailor it accordingly.

Another example comes from an information resources introduction for 1st Year Health Science students at City University myself and colleagues, Catherine Radbourne and Fiona Paterson, created in 2015.

To keep it more interesting and to combat a reduction in the time available, we created a quiz using Adobe Presenter. The aim was to let students loose with it so that they could test out CitySearch and the different resources for themselves. However, as different students were at different levels of IL and understanding of the systems being used, some finished very quickly, while others either left without finishing or stayed after the session to complete the quiz.

Although this was not deemed as a failure, it was thought that explaining the resources more at the start might have increased the understanding of participants and helped them better translate the messages in the questions.[2]

As with IR, knowledge does have an overarching presence in IL as it is about trying to instil skills sets in people; but a different sort of knowledge, understanding, is more important as a concept, in trying to reduce the noise between teacher and student.

3.3.3 Data, Information and Knowledge in Knowledge Management (KM)

The field of Knowledge Management (KM), although dismissed by some as a fad in the 1980s (Bawden 2012), deals with how organisations use, share and store tacit knowledge: that knowledge which resides in people's heads and incorporates the skills and know-how that can be difficult to express (Bawden and Robinson 2012).

As was put forward by different writers mentioned above, knowledge can be seen as being either internal or external. The battle of KM is to try and extract and store that internal knowledge and make it external, as this can be of enormous benefit to organisations.

This has been expressed by Nonaka and Takeuchi in the SECI model (see Figure 14), which charts the progress of tacit knowledge to explicit and back again. In this model, tacit knowledge can be shared by socialisation, the sort of knowledge transfer gained through person to person training such as apprenticeships. By externalising tacit knowledge through capturing and organising, it becomes explicit via publication. By combining different forms of explicit knowledge, either through computer processing or further research, new knowledge can be created. Equally, this explicit knowledge can be internalised by sharing and selection.

[2] See Atkinson and Paterson 2015 and forthcoming (Appendix B) for a presentation and article about these presentations.

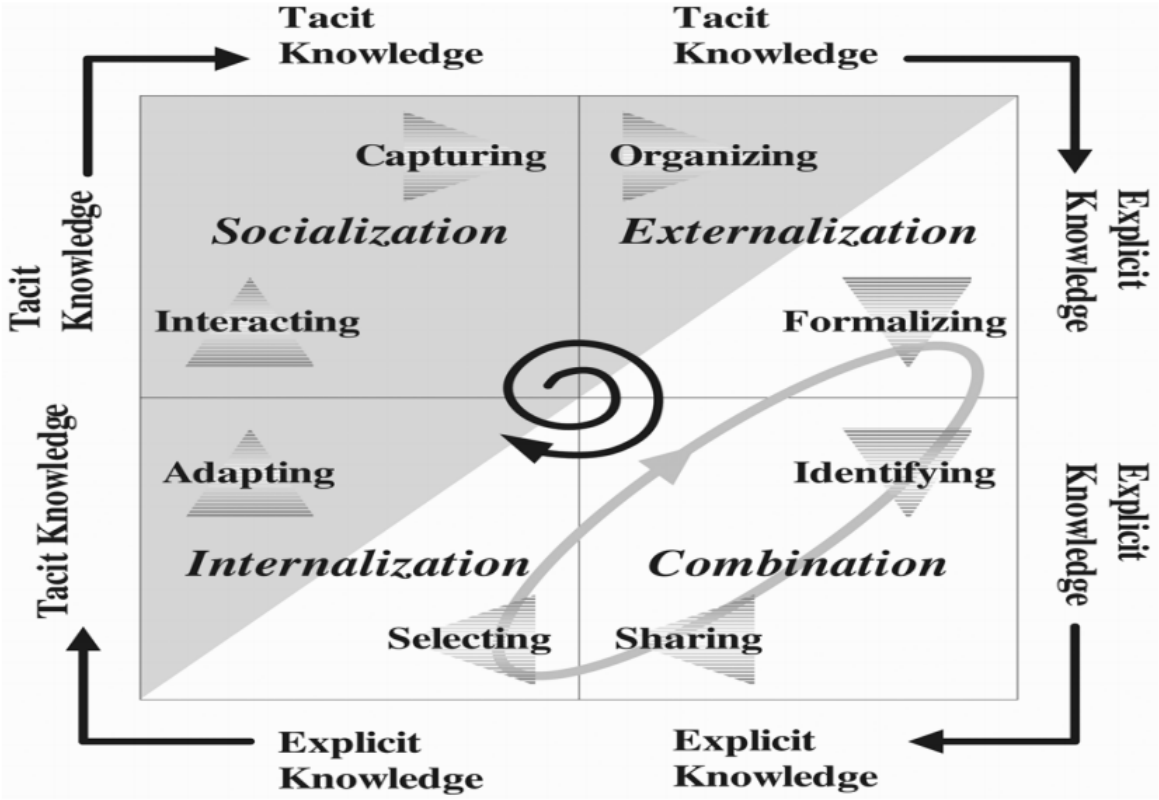

Figure 14: Nonaka and Takeuchi's SECI model (from Schenk 2014)

A similar diagram from KM (via Psychology) is the Conscious Competence Learning Matrix (Figure 15). This shows how people become aware and unaware of skills and knowledge that they have. From "not knowing how little you know" (Unconscious Incompetence), an individual will come to realise how little they know (Conscious Incompetence) before finding out what they need to know (Conscious Incompetence) and being able to perform a skill without thinking about (riding a bike, for example, without constantly thinking about your every movement as you might when first learning) (Flower 1999).

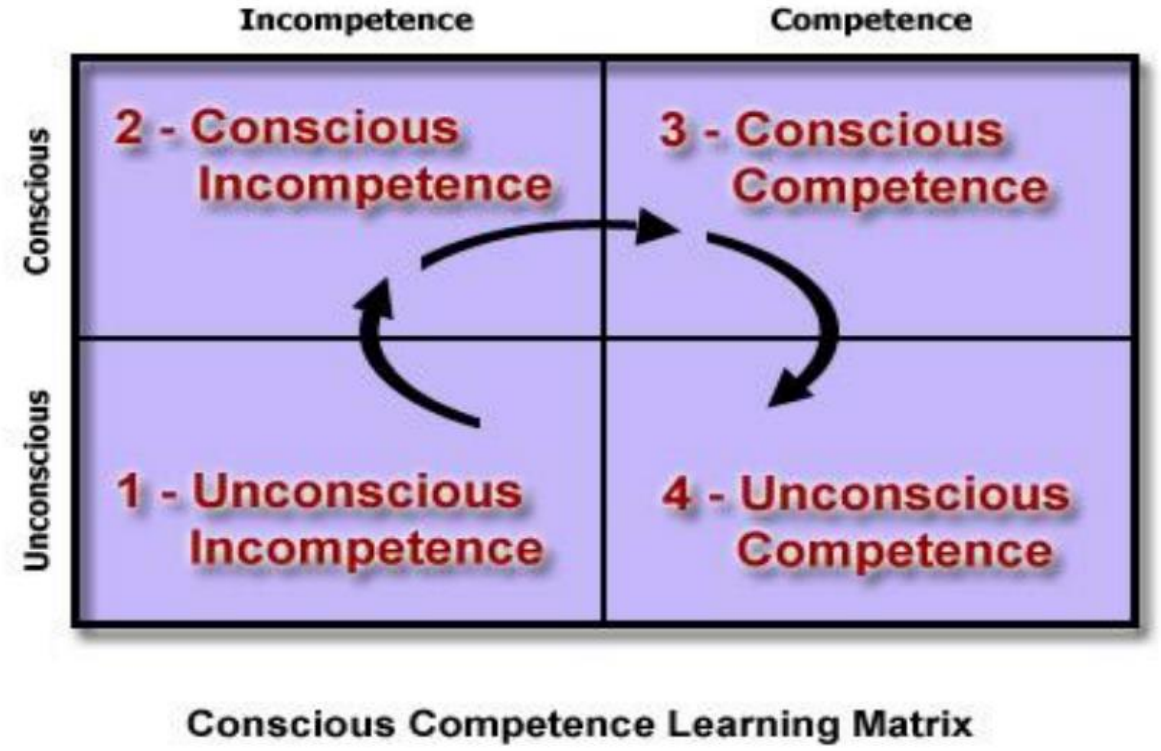

Figure 15: The Conscious Competence Learning Matrix (from Schenk 2014)

The SECI model maps the movement of knowledge from tacit to explicit and back again while the Conscious Competence Learning Matrix deals with how we learn. The former is a bit like Farradane's (1980) model (Figure 8) in that tacit and explicit knowledge are exchanged in a similar way to information being sent and received (both are also cycles, linking knowledge and information back around to the start); while the latter brings to mind ASKs and the processes of IR. In both cases, an individual would call upon data and information to complete the processes described. In KM, where the SECI model is concerned, the data and information is shared around the company as well.

KM also combines computer and human views of the relationship between data, information and knowledge by looking at ways of sharing information across systems and networks. Bawden and Robinson (2012) talk about this happening through knowledge databases that employees are encouraged to contribute to or using networks to put people in touch with one another personally. Both of which are more practical ways of looking at the relationship of data, information and knowledge that focus more on the end product and ways of disseminating it.

This final point, though, shows the importance of knowledge (through sharing tacit knowledge and explicit knowledge via documents) and possibly gives LIS a further way of looking at the data-information-knowledge relationship. Rather than simply viewing a change in the knowledge state or new knowledge in the form of documents as the end of a process, LIS should be more concerned with what happens next. Should the relationship between data, information and knowledge resemble more the cycle seen in the SECI model with the Conscious Competence Learning Matrix helping to explain some of the processes (such as learning and IR) that will happen within that cycle?

4. Reflections and Conclusions

4.1 Pullman's Dust

"Dust is only a name for what happens when matter understands itself. Matter loves matter. It seeks to know more about itself, and Dust is formed." - Balthamos in The Amber Spyglass *(Pullman 2000/2011).*

In Atkinson (2014), while looking at analogies used by Buckland (1991) and Bates (2006) for information, it was suggested that "an analogy to dust in Philip Pullman's *His Dark Materials* might be better." Although this dissertation made a slight attack on LIS writers looking too much into other disciplines, and though this is not be quite the same, I would like to look at the idea further here and suggest it as a possible analogy for the future.

In the *His Dark Materials* trilogy (consisting of *Northern Lights* (Pullman 1995/2011), *The Subtle Knife* (Pullman 1997/2011) and *The Amber Spyglass* (Pullman 2000/2011)), Dust is a an elementary particle believed at first to be original sin, but which is in fact something that interacts with humans to help them develop and create knowledge.

In the books, Dust surrounds people as they start puberty and begin their metamorphosis from child to adult. It also clings to items made by humans, for example the skulls with holes drilled in them on display in the museum in *The Subtle Knife* (Pullman 1997/2011).

Following on from the results of this dissertation, I would suggest that knowledge within LIS should be viewed more as the clouds of dust around us and our artefacts (and clogging up our brains) rather than , or as much as, the end product of new knowledge in the mind, whether seen as stores of knowledge in documents or a change in the state of someone's personal knowledge store.

The contention, as described in the statement of aims, is that Pullman's Dust forms a cloud with which we constantly interact, mostly through our own knowledge (thoughts, skill sets, facts) which we use when consulting data and information to build on our knowledge and understanding. Additionally, Dust is not a constant: we need to keep undergoing this process to learn and gather more data, information and knowledge.[3]

4.2 Reflections

[3] Other LIS references can also be seen through the books. Two characters, Lyra Belacqua and Mary Malone, each consult a kind of information retrieval device: in Lyra's case, an Alethiometer, and in Mary's, *I-Ching* sticks. For both characters, these help to fulfil their anomalous states of knowledge.

Both systems are controlled by Dust and the characters use the symbols engraved on the alethiometer and/or books to aid understanding. The Dust and aids could either be seen as replacing Metadata as the knowledge in the system or data and information that are processed to become knowledge. At the end of the books, Lyra loses the ability to use the Alethiometer, showing the importance of understanding in gaining new knowledge.

Additionally, people from Lyra's world have a daemon, an animal companion that represents their soul. This soul forms one of the three parts of a human, the other two being ghost (which lives on after your death before breaking down into Dust and rejoining the universe) and body. The daemon also forms a sort of sounding board with which a human can discuss ideas, and will even do some of the thinking and remembering for you. As such, the human and daemon in *His Dark Materials* form an outward model of what would normally be internal thought processes.

The analogy of a cloud also works neatly when considering the growing use of cloud computing in personal and professional lives. Indeed, this dissertation was written in Google Drive's cloud.

From the literature review and discussions above, it can be seen that there are a lot of differing views about the relationship between data, information and knowledge in LIS.

By looking at both LIS writers and different fields of LIS more directly, there may be a lot of evidence for trying to find a more complex and less linear view of the relationship between data, information and knowledge; and for one that incorporates knowledge in a different way.

Although seen as a product of data and information, knowledge is not always seen as the end product, as it can create both data and information, or in the case of the former, at least provide the impetus for collecting data.

The LIS profession is split into many parts, and is both theoretical and practical in nature. This split can also be seen as a focus on computer systems and a focus on humans. The fields of Information Retrieval, Information Literacy and Knowledge Management all contain theoretical and practical elements that can help influence Information Theory and the way it relates data, information and knowledge together. They also each have a focus on library and workplace computer systems, as well as the actions of people.

In all three areas, as shown in Section 4.1, knowledge plays a big part - through a mixture of understanding and learning. This knowledge is used at various stages of a process, in different ways, and by different players (humans and systems) in order to process data and information.

An additional division is made in regard to the substance of information and knowledge. Where data is generally seen to be physical, at least in its recorded form, information and knowledge can be viewed either as a physical thing or a process that occurs within the mind.

Within LIS fields such as Information Retrieval, Information Literacy and Knowledge Management, similar views along these lines can be seen: although generally knowledge is more in the mind, and information more physical.

In IR, knowledge is embedded in the system as well as in the minds of the system operators (both patrons and system librarians), while information, though an end product in various ways (the retrieval system's results and the documents gained) is also, as in IL and KM, part of the process of turning what you see and find into knowledge, as seen in Todd's representation of Brookes' ideas (Todd 1999, Figure 7)

In IL, knowledge is both seen in understanding and the skills gained through an information transfer. While in KM, knowledge can be implicit and explicit, data and information being used to gather, locate and create.

The main point to make, though, is that the relationship between data, information and knowledge is not a simple one. Each part has been given different definitions and different suggestions have been made about how they interact. The biggest difference between these is the placement and role of knowledge. And in this, it is argued that knowledge should be seen as a more central concept in how it relates to data and information. Which is why the analogy of Pullman's Dust has been suggested as a way to depict knowledge as an overarching concept with which data and information constantly interact.

4.3 Suggestion for a new model for Data, Information and Knowledge

In the Statement of Aims (Section 1.1), it was asserted that a new model of information would be suggested from the results of this dissertation.

Combining the thoughts put forward in Sections 4.1 and 4.2, an alternative model for information theory will be formalised. In the spirit of Belkin (1978, Figure 9), a list of features that need to be included will be first discussed, before presenting and explaining the model itself.

4.3.1 What does the model need to include?

According to what has already been discussed above, to be wholly satisfactory, a model that relates data, information and knowledge, needs to take several points into account.

It will need to define each of the concepts (data, information and knowledge) in order to provide the relationships in the form of a model. This is something that every writer who has formed a model has done, and will be included here in the explanation.

The model must also take into account the dual nature of various issues raised in this dissertation. It will need to account for subjective and objective views of information and knowledge; it will need to account for the various definitions of information as a thing, a process, a sign, a signal, being in a rock or the rock itself; and for knowledge being defined variously as implicit, explicit, subjective, universal, as an individual concept and one that covers or unites society. It must also cover the interaction between humans and between humans and computers.

It will need to account for the different orders of knowledge suggested by Qvortrup (1996) and the various types of knowledge that have come out of this discussion: the types of noise (as Shannon (1949/1969) named it) that affects knowledge and information such as emotion (Thellefsen et al 2014), understanding, context (Madden 2000), bias (Farradane 1979); and the parts of knowledge that assist the processing of data and information such as skills and know-how (Ackoff 1989), and the behaviours identified by Dervin (1977) as Information$_3$.

As such, the model will need to carefully consider the placement of knowledge. The placement of data and knowledge is quite standard but the LIS literature and a brief study of other fields of LIS have given various suggestions about how knowledge interacts with data and information, and thus where it should be placed within a model of their relationship.

The model will additionally need to take account of other such models whether they present themselves more or less as a linear progression (DIKW, Floridi 2010, Shannon and Weaver 1949/1969), a pyramid (DIKW, Rowley 2007), an interaction between the physical and cognitive processes (Brookes in Todd 1999), a simple equation (Shannon-Weaver 1949/1969, Brookes 1980), feedback loops (Ford 2004), the discourse between a sender and receiver of information (Farradane 1980), as a tree that grows from knowledge and returns to that earth its fruit (Buckland 2012a) or, indeed, an intricate web (Ma 2012).

In terms of this last point and Ma's web - the model will need to be more complex than some to account for everything but it would, perhaps, be ideal to be relatively simplistic in order to try and account for everything. In addition, it might be ideal to try and manage a discipline free approach as demonstrated by Losee (1997), as though a solution for LIS is being sought, if it can applied to, or by, other disciplines that cannot be bad.

However, it need not contain the truth but be as close to the truth in itself as possible.

4.3.2 A new model for Data, Information and Knowledge and explanation

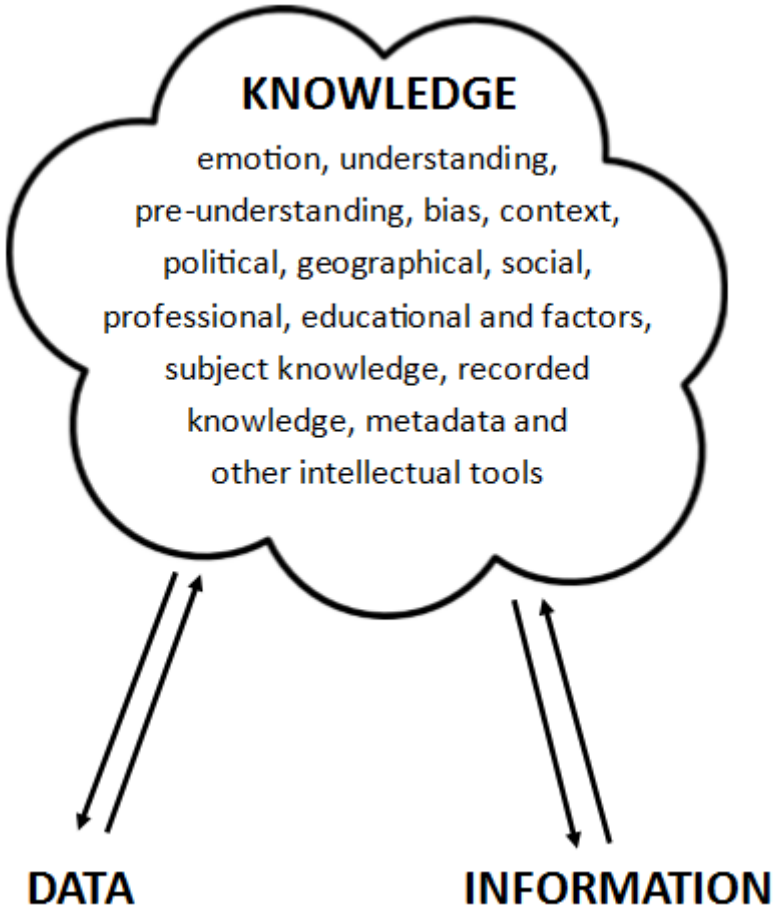

Figure 16: Suggested model for the relationship between data, information and knowledge[4]

Here, "Data" is defined as all that is gathered for a task or fed into a computer to complete one. This might be a collection of books and articles gathered in order to write an essay, the results of a scientific experiment or the sound of speech as it enters one's ears and also a search query or a dataset before it is processed.

"Information," is variously, the results after that data has been filtered through knowledge but also the results of a search query and that which is recorded in a variety of formats, such books and websites.

"Knowledge," is the filter through which data and information is passed in order to complete tasks. It is the cloud of Pullman's Dust that inspires and completes us. It is the sense that something is missing and forms an ASK, it is the understanding that processes data to become information (pre-understanding and subject knowledge) and combines pieces of information to create the end product, knowledge. It is far from perfect and can be clouded by emotion, bias and contexts such as political leanings, geographical placement, subject knowledge and other influences from society. Knowledge is also present in systems through the existence of metadata and is recorded as information in a variety of books and formats.

As such, in this model, Knowledge is seen as an overarching concept that can both generate and process data and information. Data is generated in response to an ASK and the data gathered will be processed via pre-understanding and subject knowledge to become

[4] Cloud outline taken from free online clipart website, clker.com: http://www.clker.com/clipart-cloud-outline.html

information. This information, once fully understood (or is believed to be), will form part of the mind's knowledge store (either in an individual or across society or a subject domain). This last stage is knowledge as is traditionally seen, but, as this dissertation has shown, is far more widely reaching than that.

This model, I think, takes everything listed in Section 4.3.1 into consideration. Certainly, the model would not please everybody as it seeks a catch-all approach. However, in including different terms as knowledge, or parts of and effects on it, I have sought to include various concepts that allow for knowledge to be different things and act at different parts of a process.

Nothing in this model would ever quite be standing still. Just as the mind races as it takes in different stimuli, so would a process (or several at once) be bouncing between the different parts of the model, with knowledge being continually consulted to get as close to the truth, or an end result, as possible.

Or as Elizabeth Orna, quoted in Bawden (2001), puts it: "Knowledge and information are separate but interacting entities; we transform one into the other constantly… the transformation of information into knowledge and knowledge into information, forms the basis for all human learning and communication."

Additionally, I would not say that this model is the end of the story. Within it could be transposed all the models discussed above, as could the different processes seen in different fields of LIS be discussed in relation to parts of this model. They are what would provide further detail and complexity in this relatively simple picture.[5]

An alternative could be to hold the knowledge store itself above the other components of knowledge (understanding, emotion, bias etc.) which would form a filtration system between all three concepts. My objection would be that this would not be a true reflection as the knowledge store itself sits alongside these other parts of knowledge, which is arguably shown by Thellefsen et al (2014), rather than being separate from them. In other words, even if emotion is clouding your judgement, you would still be using your knowledge store at the same time.

Either way, we see the Library (and Life) Love Triangle that is the relationship between data, information and knowledge and the elements of knowledge that bring them together.

4.4 Suggestions for Future Research

Further to this, the following suggestions for future research are to be made:

1. To review the suggestion that the Shannon-Weaver model (Figure 6) can be applied to Information Literacy.

2. To test and review the model for the relationship between data, information and knowledge suggested in Section 4.1 (Figure 16) to see if it is robust and;

3. in doing so, to explore the different types of knowledge listed. Are there, in fact, these different types of knowledge? Should the model therefore be more complex in order to take these into account, or does it work well as it is? Alternately, should the relationship be

[5] It should be noted that many of the models included above directly influenced this new model. In particular, the works of Farradane (1980, Figure 8), Todd (1999, Figure 7), and Thellefsen et al (2014, Figure 11) - all because of the way they included information and knowledge making exchanges and forming cycles; and Figure 11 for the idea of knowledge balancing emotion and information.

represented in a linear way, more like an equation with knowledge or information as the end result?

4. To further examine information theory to see if it should be left behind for LIS to focus more on the practical, or applied, theories and models explored in Section 3.3. Information theory arguably causes too much disruption and argument in the field, and focusing on how its ideas affect other areas of LIS may be a better use of this energy.

4.5 Conclusion

This dissertation sought to compare and contrast different views in LIS literature of the relationship between data, information and knowledge. In doing so, it sought to also look into different fields of LIS, other than Information Theory in order to see how different the relationship was there. It also sought to produce a new model to describe this relationship.

The findings were that there is little agreement as to how data, information and knowledge interact, and this is due to different views of the definitions of the concepts, different views of the LIS discipline and because of the widely differing views of how knowledge interacts with information.

A new model of the relationship between data, information and knowledge has been suggested (Figure 16) with future suggestions for research.

As was concluded in a previous essay on Information Theory (Atkinson 2014), this will be a never-ending debate. “An exploration of information runs into immediate difficulties,” said Buckland (1991) and the same is true of its relationship with data and knowledge. However, we can try, and we must, to get closer to the truth.

5. References and Bibliography

5.1 References

Ackoff, R.L. 1989, “From data to wisdom,” *Journal of Applied Systems Analysis*, vol. 16, no. 1, pp. 3-9

Atkinson, J. 2014, “Compare and contrast two or three viewpoints on ‘information’ from writers within the LIS domain,” INM301: Library and Information Science Foundation. Unpublished essay. City University London. Available in Appendix A.

Atkinson, J. & Paterson, F. 2015, “The (Treasure) Map to Induction: Using pirates to make a library introduction more fun,” *ALISS Xmas Event: Libraries and Innovations in Learning*, London, 16th December 2015. Available at: http://www.slideshare.net/heatherdawson/the-treasure-map-to-induction-using-pirates-to-make-a-library-introduction-more-fun (Accessed: 21st December 2015).

Atkinson, J. & Paterson, F. forthcoming, “The (Treasure) Map to Induction: Using pirates to make a library introduction more fun,” *ALISS Quarterly.* Available in Appendix B.

Bates, M.J. 2005, "Information and knowledge: an evolutionary framework for information science", *Information Research: an international electronic journal,* vol. 10, no. 4, pp. 239.

Bates, M.J. 2006, "Fundamental forms of information", *Journal of the American Society for Information Science and Technology,* vol. 57, no. 8, pp. 1033-1045.

Bates, M.J. 2008, "Hjørland's critique of bates' work on defining information", *Journal of the American Society for Information Science and Technology,* vol. 59, no. 5, pp. 842-844.

Bates, M.J., 2010, “Information”, in Bates, M.J. & Maack, M.N. (Ed.) *Encyclopedia of library and information sciences,* 3rd Ed. London, CRC Press, pp. 2347-2359.

Bawden, D. 2001, "The shifting terminologies of information", *Aslib Proceedings,* vol. 53, no. 3, pp. 93-98.

Bawden, D. 2008, "Smoother pebbles and the shoulders of giants: the developing foundations of information science", *Journal of Information Science,* vol. 34, no. 4, pp. 415-426.

Bawden, D. 2011, “Brookes equation: The basis for a qualitative characterization of information behaviours.”, *Journal of Information Science*, vol. 37, no. 11, pp. 101-108.

Bawden, D. 2012a, “On the gaining of understanding; syntheses, themes and information analysis”, *Library and Information Research*, vol. 36, no. 112, pp. Available at: http://openaccess.city.ac.uk/3169/ (Accessed: 4th January 2015).

Bawden, D. 2012b, "Fads, assimilations and knowledge management", *Journal of Documentation,* vol. 68, no. 3.

Bawden, D. 2015, "Encounter, synthesis, immersion and construction: information and the gaining of understanding", i3 Conference (Information: Interactions and Impact), Robert Gordon University, Aberdeen, 25th June 2015. Available at: http://openaccess.city.ac.uk/12150/. (Accessed: 30th September 2015).

Bawden, D. and Robinson, L. 2015a, “Information and the gaining of understanding”, *Journal of Information Science*, epub before publication. Available at: http://jis.sagepub.com/content/early/2015/12/14/0165551515621691.full.pdf+html (Accessed; 18th December 2015).

Belkin, N.J. 1975, "Some Soviet Concepts of Information for Information Science", *Journal of the American Society for Information Science (pre-1986),* vol. 26, no. 1, pp. 56-64.

Belkin, N.J., 1978, “Information concepts for Information Science”, *Journal of Documentation, vol.* 34 no. 1, pp. 55-85.

Belkin, N.J., Oddy, R.N. & Brooks, H.M. 1982, "ASK for Information Retrieval: Part I. Background and Theory" *Journal of Documentation,* vol. 38, no. 2, pp. 61-71.

Belkin, N.J. & Robertson, S.E. 1976, "Information Science and the Phenomenon of Information", *Journal of the American Society for Information Science (pre-1986),* vol. 27, no. 4, pp. 197-204.

Brookes, B.C. 1974, "Robert Fairthornes and the scope of Information Science", *Journal of Documentation,* vol. 30, no. 2, pp. 139-152.

Brookes, B.C. 1980, "The foundations of information science: Part I. Philosophical aspects", *Journal of Information Science,* vol. 2, no. 3, pp. 125-133.

Bonnevie, E. 2001, "Dretske's semantic information theory and meta-theories in library and information science", *Journal of Documentation,* vol. 57, no. 4, pp. 519-534.

Buckland, M.K. 1991, "Information as Thing", *Journal of the American Society for Information Science (1986-1998),* vol. 42, no. 5, pp. 351-360.

Buckland, M.K. 1992, "Emanuel Goldberg, Electronic Document Retrieval, and Vannevar Bush's Memex", *Journal of the American Society for Information Science (1986-1998),* vol. 43, no. 4, pp. 284-294.

Buckland, M.K. 2012a, "Interrogating Spatial Analogies Relating to Knowledge Organization: Paul Otlet and Others", *Library Trends,* vol. 61, no. 2, pp. 271-285.

Buckland, M. 2012b, "What kind of science can information science be?", *Journal of the American Society for Information Science and Technology,* vol. 63, no. 1, pp. 1-7.

Capurro, R. & Hjørland, B. 2003, “The concept of information”, in Cronin, B. (ed.) *Annual Review of Information Science and Technology,* vol. 37, pp. 343-411.

Cornelius, I.. 2002 “The concept of information”, in Cronin, B. (ed.) *Annual Review of Information Science and Technology,* vol. 36, pp. 393-425.

Dervin, B. 1977, "Useful theory for librarianship: communication, not information", *Drexel Library Quarterly,* vol. 13, no. 3, pp. 16-32.

Farradane, J. 1979, "The nature of information", *Journal of Information Science,* vol. 1, no. 1, pp. 13-17.

Farradane, J. 1980, "Knowledge, information, and information science", *Journal of Information Science,* vol. 2, no. 2, pp. 75-80.

Floridi, L. 2009, “Philosophical Conceptions of Information”, *Lecture Notes in Computer Science*, no. 5363, pp. 13-53 Available online: http://www.philosophyofinformation.net/publications/pdf/sipiat.pdf (Accessed: 5th January 2015)

Floridi L. 2010, *Information: A very short introduction.* Oxford, Oxford University Press.

Floridi, L. 2011, *The philosophy of information,* Oxford University Press.

Flower, J. 1999, "In the mush", *Physician Executive,* vol. 25, no. 1, pp. 64.

Ford, N. 2004, "Modeling cognitive processes in information seeking: From Popper to Pask", *Journal of the American Society for Information Science and Technology,* vol. 55, no. 9, pp. 769-782.

Frické, M. 2009, "The knowledge pyramid: a critique of the DIKW hierarchy", *Journal of Information Science,* vol. 35, no. 2, pp. 131-142.

Furner, J. 2002, "Shera's social epistemology recast as psychological bibliology", *Social Epistemology,* vol. 16, no. 1, pp. 5-22.

Furner, J. 2004, "Information Studies Without Information", *Library Trends,* vol. 52, no. 3, pp. 427.

Furner, J. 2015, "Information Science Is Neither", *Library Trends,* vol. 63, no. 3, pp. 362.

Gleick, J. 2011, *The information: a history, a theory, a flood,* Fourth Estate, London.

Hartel, J. 2014, "An arts- informed study of information using the draw- and- write technique", *Journal of the Association for Information Science and Technology,* vol. 65, no. 7, pp. 1349-1367.

Hjørland, B. & Nicolaisen, J. 2005/2007, “Knowledge”, *The Epistemological Lifeboat*, Available at: http://www.iva.dk/jni/lifeboat/info.asp?subjectid=91

Hjørland, B. 1998, "Theory and metatheory of information science: a new interpretation", *Journal of Documentation,* vol. 54, no. 5, pp. 606-621.

Hjørland, B. 2002, "Epistemology and the socio- cognitive perspective in information science", *Journal of the American Society for Information Science and Technology,* vol. 53, no. 4, pp. 257-270.

Hjørland, B. 2003, "Arguments for epistemology in information science: 1", *Journal of the American Society for Information Science and Technology,* vol. 54, no. 8, pp. 805.

Hjørland, B. 2007, "Information: Objective or subjective/situational?", *Journal of the American Society for Information Science and Technology,* vol. 58, no. 10, pp. 1448-1456.

Kvanvig, J.L. 2003, *The Value of Knowledge and the Pursuit of Understanding,* Cambridge University Press, Cambridge.

Losee, R.M. 1997, "A Discipline Independent Definition of Information", *Journal of the American Society for Information Science (1986-1998),* vol. 48, no. 3, pp. 254-269.

Ma, L. 2012, "Meanings of information: The assumptions and research consequences of three foundational LIS theories", *Journal of the American Society for Information Science and Technology,* vol. 63, no. 4, pp. 716.

MacFarlane, A. 2012/2013, “Lecture 04: Information Retrieval” INM348: Digital Information Technologies and Architectures, City University London. Unpublished.

Machlup, F. 1983, “Semantic quirks in studies of information”, in Machlup, F. and Mansfield, U., eds. The study of information: interdisciplinary messages.
New York, Chichester, Brisbane, Toronto, Singapore: John Wiley & Sons,
1983, 641–671

MarshalGrover 2008, *Sesame Street - Cookie and Tarah on Information*. Available at: https://www.youtube.com/watch?v=khj05iyGm1o (Accessed: 18th November 2015).

Madden, A.D. 2000, "A definition of information", *Aslib Proceedings,* vol. 52, no. 9, pp. 343-349.

McCloud, S., 1960 & Lappan, B. 1993, *Understanding comics: the invisible art,* Northampton, Mass., Kitchen Sink Press,

Nagel, J. 2014, *Knowledge: a very short introduction,* Oxford University Press, Oxford, United Kingdom.

Pullman, P., 1995/2011, *Northern lights,* Scholastic Children's Books, London.

Pullman, P., 1997/2011, *The amber spyglass,* Scholastic Children's Books, London.

Pullman, P., 2000/2011, *The subtle knife,* Scholastic Children's Books, London.

Qvortrup, Lars 2006, The Concept of “Knowledge” in the Knowledge Society and Religion as 4th Order Knowledge. ISA conference, Sociocybernetic section 2006, Durban. Partly based on "Knowledge, Education and Learning – ELearning in the Knowledge Society.” Available at: http://www.unizar.es/sociocybernetics/congresos/DURBAN/papers/qvortrup.pdf. (Accessed: 28th December 2015).

Robinson, L. & Bawden, D. 2013, “Mind the gap: transitions between concepts of information in varied domains”, in Fidelia Ibekwe-SanJuan and Thomas M Dousa (eds.) Theories of Information, Communication and Knowledge: A Multidisciplinary Approach, London, Springer

Rowley, J. 2007, "The wisdom hierarchy: representations of the DIKW hierarchy", *Journal of Information Science,* vol. 33, no. 2, pp. 163-180.

Schenk, Noeleen 2014, “Managing Tacit Knowledge,” INM341: Information Management and Policy, City University London. Unpublished.

Shannon, C.E. & Weaver, W. 1949/1969, *The mathematical theory of communication,* Illinois UP.

Thellefsen, T., Sørensen, B. & Thellefsen, M. 2014, "The information concept of Nicholas Belkin revisited – some semeiotic comments", *Journal of Documentation,* vol. 70, no. 1, pp. 74-92.

Todd, R.J. 1999, "Back to our beginnings: information utilization, Bertram Brookes and the fundamental equation of information science", *Information Processing and Management,* vol. 35, no. 6, pp. 851-870.

University of Salford c.2010, “University of Salford Information Literacy Strategy.” Avaialble at: www.salford.ac.uk/__data/assets/pdf_file/0005/264740/strategy.pdf. (Accessed: 2nd January 2016).

van den Heuvel, C. & Rayward, W.B. 2011, "Facing interfaces: Paul Otlet's visualizations of data integration", *Journal of the American Society for Information Science and Technology,* vol. 62, no. 12, pp. 2313-2326.

Whittemore, B.J. & Yovits, M.C. 1973, "A Generalized Conceptual Development for the Analysis and Flow of Information", *Journal of the American Society for Information Science (pre-1986),* vol. 24, no. 3, pp. 221.

Wilkinson, J.M. 2010, "British Library Dataset Programme: Supporting Research in the Library of the 21st Century.", *LIBER quarterly,* vol. 20, no. 1, pp. 94; 94-104; 104.

Zins C. 2004, “Knowledge organization: An epistemological perspective”, *Knowledge Organization*. vol. 31, no. 1, pp. 49-54.

Zins, C. 2006, "Redefining information science: from "information science" to "knowledge science"", *Journal of Documentation,* vol. 62, no. 4, pp. 447-461.

Zins, C. 2007a, "Conceptual approaches for defining data, information, and knowledge", *Journal of the American Society for Information Science and Technology,* vol. 58, no. 4, pp. 479-493.

5.2 Bibliography

Bates, M.J. 2011, "Birger Hjørland's Manichean misconstruction of Marcia Bates' work", *Journal of the American Society for Information Science and Technology,* vol. 62, no. 10, pp. 2038-2044.

Bawden, D. & Robinson, L. 2015b, ““A few exciting words”: Information and entropy revisited”, *Journal of the Association for Information Science and Technology*, vol. 66, no. 10, pp. 1965-1987.

Oyadonghan, J.C. 2010, "Information Flow Patterns in Organizations: The Library in Focus", *Library Philosophy and Practice,* December 2010, pp. 1.

Pettigrew, K.E. & McKechnie, L.(.F.). 2001, "The use of theory in information science research", *Journal of the American Society for Information Science and Technology,* vol. 52, no. 1, pp. 62-73.

Zins, C. 2007b, "Conceptions of information science", *Journal of the American Society for Information Science and Technology,* vol. 58, no. 3, pp. 335-350.

6. Appendices

6.1 Appendix A: Compare and contrast two or three viewpoints on 'information' from writers within the LIS domain

The following is referenced in the main body of the dissertation as Atkinson 2014 and is an essay submitted for the Library and Information Science Foundation module (INM301).

Compare and contrast two or three viewpoints on 'information' from writers within the LIS domain.

"An exploration of information runs into immediate difficulties"
(Buckland 1991)

"A conceptualization of information is obviously central to a discipline named information science."
(Bates 2006)

"It is said that we live in an Age of Information, but it is an open scandal that there is no theory, nor even definition, of information that is both broad and precise enough to make such an assertion meaningful."
(Goguen 1997 via Hjørland 2007)

These introductions to articles that will be discussed below illustrate some problems, desires and frustrations within the Library and Information Science (LIS) domain when it comes to defining the very thing we are seeking to grapple with: information. Various reviews in the literature have focused on different ways the term can be approached and defined.

From looking at the etymology and definitions of the word Capurro and Hjørland (2003) and Bawden (2001) show that even trying to use a dictionary to solve the issue is problematic- although, as I shall show, everything hinges around this.

Capurro and Hjørland (2003), along with Bawden, Bawden and Robinson (2012 and 2013), Floridi (2010) and Ma (2012) have all looked at how information is viewed outside the LIS domain and how views such as Popper's Worlds and Shannon and Wiener's theories have influenced it. Floridi (2010) and Bawden and Robinson (2013), in particular, look at how academics in domains such as biology, physics and philosophy view information; with the latter talking about the potential of drawing these ideas within those in the LIS domain- focusing on how some academics have sought to bridge the gaps between.

This essay, however, will look at three views of information from within the LIS domain, summarising each before comparing and contrasting them, drawing on points where they agree and disagree and seeking to identify the crucial differences of opinion on this matter within LIS. The writers whose viewpoints I will consider are Michael Buckland, Marcia Bates and Birgir Hjørland.

Buckland and Information-as-thing

Michael Buckland's 1991 article Information as Thing says from the outset that the term information is "ambiguous and used in different ways." In an attempt to reflect and improve on this, he proposes a threefold definition, using the Oxford English Dictionary as backup, seeing Information as a process (when it is passed from one person to another), as knowledge (when it is stored in the mind) and as thing (when it is a physical object).

As such, Buckland sees information as being tangible (when an object) and intangible (when an idea) at different times, as shown in Figure 1, with one ultimately feeding the other: the intangible becoming tangible when represented in one form or another.

	INTANGIBLE	TANGIBLE
ENTITY 2.	Information-as-knowledge Knowledge	3. Information-as-thing Data, document
PROCESS 1.	Information-as-process Becoming informed	4. Information processing Data processing

Figure 1: Buckland's Four aspects of information (Buckland 1991)

It the final of these definitions, Information-as-thing, that Buckland (1991) is most concerned with, seeking to challenge the views of Machlup (1983 via Buckland 1991) and Fairthorne (1954 via Buckland 1991) who see information as communication and the interpretation of a signal respectively. Buckland disagrees, pointing out that the expansion of information technology has caused the term information to be more widely ascribed to such things as databases and books and that information retrieval systems see information purely as a collection of things ("Libraries deal with books; computer-based information systems handle data in the form of physical bits and bytes; museums deal directly with objects").

Further discussion leads Buckland to identify different types of information: evidence (as used in science and law to expand knowledge and argue cases), data, text and documents, objects and events. Where documents are concerned, Buckland sees a variety of objects, including sound recordings, ship models and maps as documents, drawing on Paul Otlet and the documentation movement who saw documents as "any physical information resource" and, indeed, anything "potentially informative". Buckland takes this further, stating he is "*unable to say confidently of anything that it could not be information,*" using the example of a tree- which can give information of its type as well as to archaeologists who use tree rings in helping to date objects and sites.

Buckland later (1997) expanded on this while debating what makes a document with the example of Briet, who debated whether and at one points an antelope might be a document- that is isn't in the wild but becomes one once if it is taken away for study. In this, Buckland wonders if the words document and information could be synonyms.

As he initially concluded, "People are not only informed by intentional communication but by a wide variety of objects and events" (Buckland 1991) and, in doing so, puts forward his own take on what information is.

Bates and Evolutionary Psychology

Marcia Bates (2005) seeks to "find a way to think about information that effectively allows for both subjective and objective perspectives." In other words, much like Buckland, to see information as being in objects and as what forms one's knowledge. However, Bates' arguments are very different to Buckland's due to her approach via evolutionary psychology and her widening of focus to see information, seen and unseen, within everything throughout the universe.

This she achieves by using a definition of information already proposed by Edwin Parker: "*Information* is the pattern of organization of matter and energy." Bates explains that these patterns, whether seen as frost on a window or behavioural patterns, convey information that can be almost random, as in the frost, or complex, as in the cycles of activity that build and create behaviour. Further, that these patterns are the sum of parts that create something new.

In answering the question “Whose pattern of organisation?” Bates talks of the organisation throughout the universe whereby these patterns of differences can be seen in every rock, whether there are animals there to witness them or not. In other words, because the universe is not in complete chaos there are these patterns of organisation that give out information.

Bates relates this to animals by talking about how these differences surround them and are constantly absorbed, giving every living thing the information required to survive and thrive, whether done instinctively (giving meaning to a stimulus as it is detected) or in a more slow-burning way (gathering information before forming a conclusion)- a process that has developed through evolution, the individual animals best able to process this information being the ones to survive and reproduce, passing this ability on. Bates explains that this happens in patterns rather than bits of information as “detecting a pattern can be vastly more efficient in processing and storage than detecting and storing all the individual bits of information that comprise the pattern.” As brains grow more complex and take in more information, they can use short-cuts to access and use stored information. Bates uses the particular example of a shopper slowly learning about bait and switch methods used by retailers so as not to be tricked in future.

Bates ultimately combines these ideas with those of becoming informed through communication to provide two definitions of information as follows:

- Information 1: The pattern of organization of matter and energy.
- Information 2: Some pattern of organization of matter and energy given meaning by a living being.

Information 1, by being processed by a living being becomes Information 2 and can become Knowledge by “being integrated with other contents of understanding.” This, Bates says, allows for the existence of information throughout the universe as well as in the social and communicatory sense of the word - ie, “a way to think about information that effectively allows for both subjective and objective perspectives.” In application to libraries Bates books as containing Information 1, becoming Information 2 and then knowledge when read. Data Bates also breaks down into Data 1 (what is immediately around an organism, anything outside its immediate perception is Information 1) and Data 2 (information generated by humans for social purposes or new understanding). This is expanded on in Bates (2006) into a larger range of different sorts of information including that left by animals, genetic information and recorded information.

In summary, Bates sees information as being stored everywhere, existing in the realities of nature. Although not in the material itself but the patterns of organisation within.

Hjørland and Semiotics

Hjørland’s (2007) view on information was laid bare as a response to Bates. His issue is that Bates “tries to have it both ways” when arguing information can be subjective or objective. Hjørland sees this view as “not fruitful for our field,” disagreeing that there is even a need for an objective view of information.

Early in his argument, Hjørland says there is a division in viewpoints on information that pitches the Objective view (any difference is information) vs the Subjective/Situtional (information is a difference that makes a difference).

Hjørland falls in the latter camp: using semiotics, he goes on to say that information is a signal that has been received. In other words that information exists only when the process of transferring that signal from one person to another has been completed. And thus Hjørland disagrees with Bates' dual information definitions Information 1 and 2. In discussing Information 2, Hjørland points out that it is not only for living animals that this definition is required but computers, which use code subjectively created to process and create information.

In support of this, Hjørland uses an example from Karpatschof's (2000 via Hjørland 2007) article about information. He does so through the concept of release mechanisms: a system where appropriate signals cause the physical release of mechanisms- the information appearing upon the completion of the process. This explanation, says Hjørland, allows for a definition of information that covers living beings and computers.

In criticising the evolutionary framework used by Bates, Hjørland says the reaction of animals is to signs not information, using Leontyev's levels of biological development that give five levels of creatures able to detect and decipher more types of signals.

In reference to Buckland (1991) also, Hjørland continues by criticising the Information-as-thing approach as reducing the meaning of information to nothing if it is everything (something Buckland admits without concern and Bates sees as only occurring where animals are concerned). And here Hjørland gets to the crux of his point - for something to be information it must answer a question. Nothing is in itself informative but must have knowledge and questions applied to gain that information.

Comparisons, Contrasts and Conclusions

Above we see three views on information that overlap in places but also disagree to a large extent. For example, Buckland and Bates both talk about information in objects. However, where Buckland is not entirely sure that everything can give information, Bates is certain of it. They also come at the issue from very different angles. Buckland seems only really to approach from a purely LIS point of view, seeing information in documents and data of various sorts, from books and articles to items used as evidence in science- particularly through Briet.

Bates, however, via evolutionary psychology, goes much deeper. She, rather than just looking at what she calls Recorded information ("Communicatory or memorial information preserved in a durable medium" Bates 2006) and Information 2 to discuss how living beings react to information in the world around them, often unconsciously and instinctively. Where perhaps Buckland might see pools of information (he uses the term "repository"), Bates sees a whole sea of information covering the world, and beyond (though an analogy to dust in Philip Pullman's *His Dark Materials* might be better).

Hjørland, however, sees information as the completion of a process that results in someone becoming informed. Something Bates and Buckland do agree upon but believe information stretches further.

In coming to these conclusions, the three use similar arguments and examples, albeit drawing different conclusions. For example, Buckland and Hjørland both use computers to create different views on information. Buckland uses the example of information retrieval systems that give physical objects as resources while Hjørland talks of the code running computers as the signals that are processed to create information. Bates, and Hjørland in response, draw on academics who have also written on information and evolution, each finding one that helps support their own point of view.

Additionally, there are similarities and agreed upon points shared among them, for example seeing information as the process of being informed. As Hjørland points out the crux of the divisions is whether or not they believe information can be objective. This ultimately pits various LIS writers against one another and that everything seems to hinge around the point at which you believe information becomes information. Buckland and Bates see objects, from trees and rocks to books and articles as information yet Hjørland disagrees, seeing information coming into being later, when the data is used to inform. So, can information be a thing or is it a signal?

In this, one could argue that it is all a matter of definitions of language. Bawden (2001) may well hit the nail on the head when he likens the arguments to Wittgenstein's language games. He also quotes Liebnau and Backhouse, who remarked that "Numerous definitions have been proposed for the term 'information', and most of them serve well the narrow interests of those defining it". A cynical point of view, perhaps, but worth investigating.

As mentioned above, Bawden and Capurro and Hjørland have all used the dictionary as an introduction to this issue. The latter do so in the most detail and in a most interesting way, tracing the etymology of the word information right back to its Greek and Latin roots.

These earliest roots of the word are various including giving form to something, an educator (informer), things impressed on our souls by the Gods, visual and internal perceptions. In the Middle Ages *informatio* was "the sense of providing something with a form," in particular, moulding the mind or matter (in higher and lower senses of the word). This higher sense was further developed by philosophers before the nineteenth century into becoming "communicating something to someone." In particular, things were said to inform the senses. It is only in the twentieth century that other ideas about what information is started to enter and widen definitions via the development of Information Theory, in particular through the work of Shannon and Weaver.

The ideas that Capurro and Hjørland (2003) present, however, all relate to information as signals or messages, never as matter; neatly explaining Hjørland's position and begging the question. The idea Bates and Buckland put forward, of information as matter, is presented by Capurro and Hjørland as an obsolete medieval definition. However, even if this is the case, it is an idea that is widespread.

What this etymology shows, though, is that you could easily use any of these definitions to present your own view of information- Bates' evolutionary ideas, in particular the ability to instinctively process information in nature being hardwired into us could be said to be a modern form of the things impressed on our souls root. It might be a bit cynical but this argument seems to me to about preference. In this, there is nothing wrong or surprising. Etymologies of words show their evolution, something that can often be charted within one's own lifetime - why should the meaning of information not continue to evolve? Especially since the inception of Library and Information Science as an academic field.

Really, though, the matter is far more complex and to make that conclusion would highlight a lack of understanding. Perhaps we should focus on what none of these viewpoints do. Bawden and Robinson (2012) note that Buckland does not provide "a precise or formal description of information" and Bates "introduces a potentially confusing multiplicity of forms of information." Hjørland, too, does not provide a formal definition but instead explains his point of view in opposition to another.

Additionally, none quite fully look at information as a process. They describe at length the changes that create information but do not lay it out as the DIKW model (Bawden and Robinson 2012) or Floridi (2010) do. Both of these examples provide a clearer flow of where

information comes from and where it goes. In themselves they may not be perfect but are perhaps better than Buckland's three definitions, Bates' lists of information types and Hjørland's simple definition. Both not only allow for the different types of information discussed above but lay out a firmer definition and separation of data and documents that I don't think Buckland, Bates or Hjørland are concerned enough with. There is a further argument here about what documents are that I think it would be more helpful for Buckland and Bates to be contributing to. In learning about this at City University London I attended a series of lectures called The Story of Documents (as opposed to information, although that did feature). Perhaps it is documents that are everywhere rather than information. Untapped information, perhaps, but in the form of documents.

I think Buckland and Bates both inadvertently undermine their own arguments at times as they talk of how objects need someone to see the evidence or extract the information- surely, therefore, these objects are documents from which the information is gleaned. Having said that, as a library information assistant it is difficult not to see the books on the shelves as containing information.

Overall, it is clear that this is a never ending debate that the meaning of information continues to evolve as it has done for over 2500 years at least. Each viewpoint is different and has its own nuances. "An exploration of information runs into immediate difficulties," said Buckland (1991). These difficulties may never end.

Word Count: 2920

Bibliography

Bates, M.J. (2005) Information and knowledge: an evolutionary framework for information, *Information Research* 10(4)

Bates, M.J. (2006) Fundamental forms of information, *Journal of the American Society for Information Science and Technology* 57 (8) 1033-45

Bawden, D. (2001) The shifting terminologies of information, *Aslib Proceedings* 53 (3) 93-8

Bawden, D. and Robinson, L. (2012) Introduction to Information Science, London:Facet

Bawden, D. and Robinson, L. (2013) Mind the gap: transitions between concepts of information in varied domains in Fidelia Ibekwe-SanJuan and Thomas M Dousa (eds.) Theories of Information, Communication and Knowledge: A Multidisciplinary Approach, London:Springer

Buckland, M. (1991) Information as thing, *Journal of the American Society for Information Science* 42 (5) 351-60

Buckland, M. (1997) What is a 'document'?, *Journal of the American Society for Information Science* 48 (9) 804-9

Capurro, R. and Hjørland, B. (2003) The concept of information, *Annual Review of Information Science and Technology* 37 343-411

Hjørland, B. (2007) Information: objective or subjective/situational?, *Journal of the American Society for Information Science and Technology* 58 (10) 1448-56

Floridi, L. (2010) Information: a very short introduction, Oxford:Oxford University Press

Ma, L. (2012) Meanings of information: the assumptions and research consequences of three foundational LIS theories, *Journal of the American Society for Information Science and Technology* 63 (4) 716-23

6.2 Appendix B: The (Treasure) Map to Induction: Using pirates to make a library introduction more fun

The following is referenced in the main body of the dissertation as Atkinson and Paterson (forthcoming) and is an article that is currently unpublished but that has been accepted for publication in the ALISS (Association of Librarians and Information Professionals in the Social Sciences) Quarterly newsletter in 2016.

The (Treasure) Map to Induction: Using pirates to make a library introduction more fun- Fiona Paterson and James Atkinson

Fiona.Paterson.2@city.ac.uk James.Atkinson.1@city.ac.uk @JamesAtkinson81

Introduction

During summer 2015, City University London's Nursing and Midwifery Subject Librarian, Catherine Radbourne, set out to make a resources introduction for 1st Year students more exciting. Traditionally a librarian would stand at the front of a room and talk through resources while demonstrating them onscreen. Ideally, she wanted to give students a more interactive and interesting experience. The drive for a change came partly from hearing and reading about different creative teaching techniques and wanting to try these out; but also because the time for the workshop had been cut by 50%. As fellow members of the Library's Information Literacy Group, we were drafted in and the project began.

Planning

The session would follow on from a general library induction and needed to be easy to replicate as various staff members would deliver it. It needed to cover a range of resources such as print and e-books, e-journals and specialist resources including Visible Body, Elsevier Clinical Skills and BMJ Best Practice. Additionally, as it has its own special login process, each session had to start with the attendees setting up Elsevier Clinical Skills accounts.

Once the parameters of the session were established, we had a brainstorm of ways to make it fun. The ideas we came up with included:

- Bingo
- Spot the difference
- Matching games
- Storytelling
- Case studies
- Metaphors
- Quizzes

An initial quiz idea developed into a treasure hunt so that it could include a storytelling element as well.

We then started to think about what software could support us best. Adobe Presenter was tested and alternatives such as Prezi were explored. Adobe Presenter seemed to offer what we required: the ability to make an online, interactive quiz that would allow the presenter to circulate and answer questions.

We felt a video would be good to introduce the treasure hunt, setting the scene and thus the story. We decided on a pirate theme, as it was a treasure hunt, but we also needed a topic. We considered diabetes and heart disease, as these are topics that come up when students undertake local community projects and went with the latter as there were associated procedures available on Elsevier Clinical Skills.

Production

Quite early on we started work on the videos: one to introduce the theme and set the task, one to conclude. Catherine wrote a draft script which we helped to hone and she also made contact with colleagues in the Learning Enchancement and Development (LEaD) team who were keen to collaborate and try something a bit different. Our contact volunteered to make a short animation for the beginning. We found our location (the Ye Olde Mitre in Ely Place) after Catherine embarked on an historic pub walk. We were fortunate enough to get the funding required and started to plan the filming itself.

Simultaneously we were designing and building the quiz too. We knew what resources and questions we wanted, and, as Presenter is setup with various different sorts of questions such as true/false and multiple choice, we sought to use a variety of question types throughout to help add to the experience.

We also developed a hand out that students would use to help complete the quiz. It was designed as a treasure map that, on one side featured a map with images representing the different resources while the reverse featured guidance on finding the correct answers.

Trials and Filming

With the quiz and hand out written we tested them on a group of willing librarians. Mostly this went well but we found the “pirate” font we had used in the quiz proved too difficult to read and our volunteers felt the wording of some of the questions needed a bit more clarity.

We also discovered the e-book we had chosen only supported a small number of concurrent users and so we had to change it. They felt, however, that we had a good range of difficulty in the questions and liked that it would get students navigating different resources.

With script and storyboard prepared by Catherine, costumes sourced, lines rehearsed, shooting took place early the morning of a tube strike! The filming went well and we nervously awaited the results, which were really good.

Our second trial came the day before we ran the sessions and was the first time we had incorporated the video. Everything ran smoothly and we got good feedback. This gave us a lot of confidence for the following day. However, disaster struck when Visible Body stopped working in the afternoon. Fortunately, after liaising with colleagues, we were able to implement a workaround.

Delivery and Reflection

On 17th September 2015, three months after the initial meeting, four people delivered a total of 12 sessions with 3 occurring simultaneously across 4 slots. Each session had one

presenter, lasted an hour and took place in an IT room. The running order of each session was as follows:

- Introduction and signing up to Elsevier Clinical Skills
- Introduction video
- Quiz, including prize giving
- Conclusion video
- Gathering of feedback

Reflecting on the experience, there were a few areas that provided challenges and could be looked at for improvement. The Visible Body workaround was semi successful. Resources can break at any time and is difficult to prevent but it highlighted the need for a back-up plan.

Setting up Elsevier Clinical Skills accounts is an awkward process made worse when 40 students all try at once. A lot of students needed individual assistance and we overwhelmed the system causing confirmation emails to be delayed. We're looking at the possibility of bulk uploading student data in future, although we would still need to go through the log in process which is more complex than with other resources.

The difficulty of helping 40 students at once wasn't restricted to setting up Elsevier Clinical Skills accounts and the fact that students worked at their own pace meant it was hard to bring everyone together to highlight particular issues. This was one of the reasons that the Visible Body work around was only semi successful. If the sessions were done again, we would either need to present to smaller groups or have extra staff on hand to help answer queries.

We received completed feedback forms from 109 of the 400 attendees. Feedback was good (64 respondents thought the session was good, 23 excellent, 18 fair and 4 poor) although comments were mixed. The issues highlighted in the comments were similar to the challenges we had found ourselves, namely setting up Elsevier Clinical Skills accounts, technical difficulties and lack of presenters. 95% thought the session length was just right but from experience we found there was a large range of speeds in which people completed the quiz, with some students leaving early and others staying behind. There were some comments on the clarity of the instructions. We didn't feel the hand outs were utilised in the way we'd hoped. We really wanted the quiz to be self-led but because the resources were new, it may be that more guidance was necessary in the form of demonstration. A small percentage thought the use of the theme was childish but with the range of communication and learning styles across a group this is inevitable.

There are plans to run the sessions again but with a different theme. Other members of Academic Services also thought Adobe Presenter had great potential for creating fun and interactive introductions to resources.

6.3 Appendix C: Dissertation Proposal

Dissertation Proposal
James Atkinson

Working Title

A Library Love Triangle? : An analysis of the relationship between data, information and knowledge

Introduction

Across the decades, many writers have looked tried to define and discuss issues surrounding data, information and knowledge within the Library and Information Science (LIS) sector, often drawing on other disciplines. What is lacking, however is any kind of consensus as different ideas and opinions are merely presented and argued.

In an earlier essay for this course, the writer looked at the work of three LIS writers, Buckland, Bates and Hjørland, in seeking to define information. No consensus was found and this dissertation proposes to expand that essay and examine further the relationship between data, information and knowledge as presented in LIS literature and cross examine these information theories with ideas from other areas within the LIS domain, such as knowledge management, information retrieval and information literacy.

Aims and objectives

The aim of the dissertation is:

- To comprehensively review and discuss the Library and Information Science literature on the relationship between data, information and knowledge

The objectives of the dissertation are:

- To try and find consensus and points of agreement or indeed disagreement within this area of research
- To investigate if there is a better way to explain the issues through the study of other Library and Information Science research areas in order to try and find a common ground for consensus
- To further explore an idea from an earlier essay that Pullman's idea of Dust from the *His Dark Materials* books may form a better way of looking at this issue
- And, following on from the last point, to investigate the author's contention that knowledge should be more involved in information theory, that it both precedes and follows information; and potentially to suggest a new model of information

Scope and definition

Information theory has been studied since the Second World War in different ways by different researchers, sometimes in relation to knowledge and data, sometimes set apart. Shannon, for example, viewed information purely as communications; Buckland saw it as a physical thing; Bates in relation to evolutionary psychology; Hjørland as signals via semiotics (Bawden and Robinson 2012). Added to this, researchers have also approached information from a philosophical point of view while others have used models from areas

such Biology and Physics - Floridi (2010 and 2011) looks at all these areas while proposing a philosophy of information.

Researchers within Library and Information Science have also made approaches to knowledge, such as the two opposing models discussed by Bawden and Robinson (2012): an approach adapted from Popper's Worlds and the D-I-K-W pyramid model from Zin.

Elsewhere in LIS, different models involving information and knowledge have been proposed separately from Information Theory, for example, in Information Literacy, Information Retrieval and Knowledge Management. These all show interactions between information and knowledge that could be key to creating a practical understanding of their relationship. It is from this that the writer proposes to introduce the concept of Dust, which will be further explored in the literature review below.

The definitions of information and knowledge will be fully explored in the dissertation, as the latter was already in a previous essay. Bawden and Robinson (2012) and Capurro and Hjørland (2007), for example, each give interesting accounts which will be explored.

Research context/literature review

Capurro and Hjørland, along with Bawden, Bawden and Robinson (2011 and 2014), Floridi (2011) and Ma (2012) have all looked at how information is viewed outside the LIS domain and how views such as Popper's Worlds and Shannon and Wiener's theories have influenced it. Floridi (2011) and Bawden and Robinson (2014), in particular, look at how academics in domains such as biology, physics and philosophy view information; with the latter talking about the potential of drawing these ideas within those in the LIS domain-focusing on how some academics have sought to bridge the gaps between.

Within LIS, there are various views on the topic. Michael Buckland's 1991 article Information as Thing sees information as being tangible (when an object) and intangible (when an idea) at different times, as shown in Figure 1, with one ultimately feeding the other: the intangible becoming tangible when represented in one form or another.

It is with Information-as-thing that Buckland (1991) is most concerned with, seeking to challenge the views of Machlup (1983) and Fairthorne (1954) who see information as communication and the interpretation of a signal respectively. Buckland disagrees, pointing out that the expansion of information technology has caused the term information to be more widely ascribed to such things as databases and books and that information retrieval systems see information purely as a collection of things ("Libraries deal with books; computer-based information systems handle data in the form of physical bits and bytes; museums deal directly with objects").

As he initially concluded, "People are not only informed by intentional communication but by a wide variety of objects and events" (Buckland 1991) and, in doing so, puts forward his own take on what information is.

Marcia Bates (2005) seeks to "find a way to think about information that effectively allows for both subjective and objective perspectives." In other words, much like Buckland, to see information as being in objects and as what forms one's knowledge. However, Bates' arguments are very different to Buckland's due to her approach via evolutionary psychology and her widening of focus to see information, seen and unseen, within everything throughout the universe.

Bates relates this to animals by talking about how these differences surround them and are constantly absorbed, giving every living thing the information required to survive and thrive,

whether done instinctively (giving meaning to a stimulus as its is detected) or in a more slow-burning way (gathering information before forming a conclusion)- a process that has developed through evolution, the individual animals best able to process this information being the ones to survive and reproduce, passing this ability on. Bates explains that this happens in patterns rather than bits of information as "detecting a pattern can be vastly more efficient in processing and storage than detecting and storing all the individual bits of information that comprise the pattern." As brains grow more complex and take in more information, they can use short-cuts to access and use stored information. Bates uses the particular example of a shopper slowly learning about bait and switch methods used by retailers so as not to be tricked in future.

Bates ultimately combines these ideas with those of becoming informed through communication to provide two definitions of information as follows:

- Information 1: The pattern of organization of matter and energy.
- Information 2: Some pattern of organization of matter and energy given meaning by a living being.

Information 1, by being processed by a living being becomes Information 2 and can become Knowledge by "being integrated with other contents of understanding." This, Bates says, allows for the existence of information throughout the universe as well as in the social and communicatory sense of the word - ie, "a way to think about information that effectively allows for both subjective and objective perspectives." In application to libraries Bates books as containing Information 1, becoming Information 2 and then knowledge when read. Data Bates also breaks down into Data 1 (what is immediately around an organism, anything outside its immediate perception is Information 1) and Data 2 (information generated by humans for social purposes or new understanding). This is expanded on in Bates (2006) into a larger range of different sorts of information including that left by animals, genetic information and recorded information.

In summary, Bates sees information as being stored everywhere, existing in the realities of nature. Although not in the material itself but the patterns of organisation within.

Hjørland's (2007) view on information was laid bare as a response to Bates. Early in his argument, Hjørland says there is a division in viewpoints on information that pitches the Objective view (any difference is information) vs the Subjective/Situtional (information is a difference that makes a difference).

Hjørland falls in the latter camp: using semiotics, he goes on to say that information is a signal that has been received. In other words that information exists only when the process of transferring that signal from one person to another has been completed. And thus Hjørland disagrees with Bates' dual information definitions Information 1 and 2. In discussing Information 2, Hjørland points out that it is not only for living animals that this definition is required but computers, which use code subjectively created to process and create information.

In reference to Buckland (1991) also, Hjørland continues by criticising the Information-as-thing approach as reducing the meaning of information to nothing if it is everything (something Buckland admits without concern and Bates sees as only occurring where animals are concerned). And here Hjørland gets to the crux of his point - for something to be information it must answer a question. Nothing is in itself informative but must have knowledge and questions applied to gain that information.

Bawden and Robinson note that Buckland does not provide "a precise or formal description of information" and Bates "introduces a potentially confusing multiplicity of forms of

information." Hjørland, too, does not provide a formal definition but instead explains his point of view in opposition to another.

Additionally, none quite fully look at information as a process. They describe at length the changes that create information but do not lay it out as the DIKW model (Bawden and Robinson 2012) or Floridi (2011) do. Both of these examples provide a clearer flow of where information comes from and where it goes. In themselves they may not be perfect but are perhaps better than Buckland's three definitions, Bates' lists of information types and Hjørland's simple definition. Both not only allow for the different types of information discussed above but lay out a firmer definition and separation of data and documents that I don't think Buckland, Bates or Hjørland are concerned enough with.

This is just the start of the argument. As Bawden and Robinson show across two chapters, there are various ideas about information and knowledge that will need to be explored and a decision will need to be made on how much can be realistically discussed within the dissertation itself.

In information retrieval, Belkin et al (1982) explain that the information needs of users can be explained by the gaps in their knowledge they are trying to fill, or their Anomalous States of Knowledge (ASK); and that information systems must be based on this ASK. Information literacy similarly looks at the filling of gaps of knowledge through the gaining of knowledge in the form of literacy skills. In other words, knowledge is required to retrieve and understand information.

Again, within knowledge management, interactions between knowledge and information are discussed via the ideas of Nonaka and Takeuchi and Learning Theory, which look at the sharing and gaining of information and knowledge.

Arguably, therefore, information and knowledge are intertwined, each feeding off one another; or perhaps knowledge is overarching, maybe even surrounding information. Which is where the idea of Dust comes in - when discussing information, Buckland might see pools of information (he uses the term "repositories"), while Bates sees a whole sea of information covering the world, and beyond. Perhaps, an analogy to the dust in Philip Pullman's *His Dark Materials* (2011a, 2001b and 2011c) might be better. Dust there, however, is representative of knowledge. There are those who see information in everything and this dissertation will seek to show knowledge is the dust floating between.

Elsewhere in LIS, however, three concepts will be reviewed to see if they can be of use of the arguments concerning information and knowledge

Methodology

The dissertation will be entirely desk-based, consisting of literature reviews and discussions following on from their findings. That is, reviews of LIS research into the nature of information and knowledge and their relationship followed by reviews of ideas from other relevant areas of LIS research and discussion following-on seeking to meet the objectives listed above.

Work plan

Activity	Date(s) to be achieved
Dissertation Proposal and Submission	May 2015
Preparing and Undertaking Literature Review	May to August 2015
Consolidating Knowledge and Preparing to Write Dissertation	August to Mid September 2015
Writing Dissertation	Mid September to Early December 2015
Proof Reading	December 2015
Submission	January 2016

I will try to submit in very early January, approximately one week before the deadline, in case of last minute issues and will endeavour to meet with my supervisor in June, August and October at least to discuss progress and any problems I might be facing.

Resources

As the dissertation will be entirely desk-based, the resources required will be those of the desk researcher. Literature reviews will be conducted using LISTA and Web of Science via accounts available through City University London with the main aim of finding articles. Further searches will be made via City Library Search to find further articles and also books. Initial reading should lead to further sources being found. These will either be found via City University or other libraries, such as the British Library and other London university libraries. Union catalogues such as copac and the British Library catalogue will be used should City University not have the required resources.

Ethics and Confidentiality

Research Ethics Checklist:

If the answer to any of the following questions (1 – 3) is NO, your project needs to be modified.

1. Does your project pose only minimal and predictable risk to you (the student)? Yes

2. Does your project pose only minimal and predictable risk to other people affected by or participating in the project? Yes

3. Is your project supervised by a member of academic staff of the School of Informatics or another individual approved by the module leaders? Yes

If the answer to either of the following questions (4 – 5) is YES, you MUST apply to the University Research Ethics Committee for approval. (You should seek advice about this from your project supervisor at an early stage.)

4. Does your project involve animals? No

5. Does your project involve pregnant women or women in labour? No

If the answer to the following question (6) is YES, you MUST complete the remainder of this form (7 – 19). If the answer is NO, you are finished.

6. Does your project involve human participants? For example, as interviewees, respondents to a questionnaire or participants in evaluation or testing? No

Bibliography and Suggested Initial Reading for Dissertation

Accardi, M.T., Drabinski, E. and Kumbier, A. (2010) Critical library instruction: theories and methods, Library Juice Press, Duluth, MN.

Bawden, D. (2007) Information Seeking and Information Retrieval: The Core of the Information Curriculum?, *Journal of Education for Library and Information Science*, 48(2).

Bawden, D. and Robinson, L. (2012) Introduction to information science, Facet, London.

Bawden, D. and Robinson, L. (2014) Mind the Gap: Transitions Between Concepts of Information in Varied Domains in Ibekwe-San Juan, F. and Dousa, T.M. (2014) Theories of Information, Communication and Knowledge: A Multidisciplinary Approach, Springer Netherlands, Dordrecht, 121-141.

Bates, M.J. (2006) Fundamental forms of information *Journal of the American Society for Information Science and Technology* 57(8) 1033-1045.

Bates, M.J. (2008) Hjørland's critique of bates' work on defining information *Journal of the American Society for Information Science and Technology* 59(5) 842-844.

Bates, M.J. (2011) Birger Hjørland's Manichean misconstruction of Marcia Bates' work, *Journal of the American Society for Information Science and Technology*, 62(10) 2038-2044.

Belkin, N.J., Oddy, R.N. and Brooks, H.M. (1982) ASK for information retrieval: part 1: Background and theory *Journal of Documentation* 38(2) 61-71

Buckland, M.K. 1991, Information as thing *Journal of the American Society for Information Science* 4(5) 351-360.

Buckland, M.K. (1991) Rejoinder *Journal of the American Society for Information Science* 42(10) 758.

Buckland, M.K. (1995) The Centenary of "Madame Documentation: Suzanne Briet, 1894-1989 *American Society for Information Science Journal* 46(3) 235-237.

Buckland, M.K. (2012) Interrogating Spatial Analogies Relating to Knowledge Organization: Paul Otlet and Others *Library Trends* 61(2) 271-285.

Capurro, R. & Hjørland, B. (2003) The concept of information Annual Review of Information Science and Technology 37(1) 343-411.

Clement, T. (2014) The Ear and the Shunting Yard: Meaning Making as Resonance in Early Information Theory *Information & Culture* 49(4) 401-426.

Foster, A. and Rafferty, P. (2011) Innovations in information retrieval: perspectives for theory and practice, Facet, London.

Fisher, K.E., Erdelez, S., and McKechnie, L. (2005) Theories of information behavior, Published for the American Society for Information Science and Technology by Information Today, Medford, N.J.

Floridi, L. (2010) Information: a very short introduction, Oxford University Press, Oxford.

Floridi, L. (2011) The philosophy of information, Oxford University Press, Oxford. Link

Gleick, J. (2011) The information: a history, a theory, a flood, Fourth Estate, London.

Godwin, P. & Parker, J. 2008, Information literacy meets Library 2.0, Facet, London.

Gray, R.M. (2011) Entropy and Information Theory, Springer US, Boston, MA.

Grobler, H., Du Toit, D. & Schenck, R. (2003) Person-centred communication: theory and practice, Oxford University Press, Oxford.

Hjørland, B. (1997) Information seeking and subject representation: an activity-theoretical approach to information science, Greenwood Press, Westport, Conn; London.

Hjørland, B. (1998) Theory and metatheory of information science: a new interpretation *Journal of Documentation* 54(5) 606-621.

Hjørland, B. (2000) Documents, memory institutions and information science *Journal of Documentation* 56(1, pp. 27-41.

Hjørland, B. (2002) Epistemology and the socio- cognitive perspective in information science *Journal of the American Society for Information Science and Technology* 53(4) 257-270.

Hjørland, B. (2003) Arguments for epistemology in information science: 1 *Journal of the American Society for Information Science and Technology* 54(8) 805.

Hjorland, B. (2003) Arguments for epistemology in information science: 2 *Journal of the American Society for Information Science and Technology* 54(8) 805.

Hjørland, B. (2004) Arguments for philosophical realism in library and information science, *Library Trends* 52(3) 488-506+668.

Hjørland, B. (2007) Information: Objective or subjective/situational? *Journal of the American Society for Information Science and Technology* 58(10) 1448-1456.

Hjorland, B. (2005) Library and information science and the philosophy of science *Journal of Documentation* 61(1) 5-10.

Hjørland, B. (2009) The controversy over the concept of "information": A rejoinder to Professor Bates *Journal of the American Society for Information Science and Technology* 60(3) 643-643.

Hjørland, B. (2010; 2009) The foundation of the concept of relevance *Journal of the American Society for Information Science and Technology* 61(2) 217-237.

Hjørland, B. (2011) Theoretical clarity is not ‘Manicheanism’: A reply to Marcia Bates, *Journal of Information Science*, 37 (5) 546-550.

Hjørland, B. (2011) The importance of theories of knowledge: Indexing and information retrieval as an example *Journal of the American Society for Information Science and Technology* 62(1) 72-77.

Hjørland, B. (2011) The importance of theories of knowledge: Browsing as an example *Journal of the American Society for Information Science and Technology* 62(3) 594-603.

Hjørland, B. (2013) Theories of knowledge organization - Theories of knowledge *Knowledge Organization* 40(3) 169-181.

Ibekwe-San Juan, F. and Dousa, T.M. (2014) Theories of Information, Communication and Knowledge: A Multidisciplinary Approach, Springer Netherlands, Dordrecht.

IEEE/IET Electronic Library (IEL) Proceedings By Volume & Institute of Electrical and Electronics Engineers,Inc.Staff (2000) Information Theory: 2000 IEEE International Symposium On, I E E E, Piscataway.

Jaeger, P.T., and Burnett, G. (2009) Information worlds: social context, technology, and information behavior in the age of the internet, Routledge, Abingdon; New York, N.Y.

Latham, K.F. (2012) Museum object as document: Using Buckland's information concepts to understand museum experiences *Journal of Documentation* 68(1) 45-71.

Leckie, G.J., Given, L.M. & Buschman, J. (2010) Critical theory for library and information science: exploring the social from across the disciplines, Libraries Unlimited, Oxford.

Lloyd, A., Talja, S. & Elsevier Freedom Collection EBA (2010) Practising information literacy: bringing theories of learning, practice and information literacy together, Centre for Information Studies, Wagga Wagga, Australia.

Nagel, J. (2014) Knowledge: a very short introduction, Oxford University Press, Oxford, United Kingdom.

Ozog, S. (1979) The Logics and Boundary Concepts of Information Acquisition *Journal Of The American Society For Information Science* 30(6) 364-367.

Pullman, P., (2011) Northern lights, Scholastic Children's Books, London.

Pullman, P., (2011) The subtle knife, Scholastic Children's Books, London.

Spink, A. & Heinström, J., (2011) New directions in information behaviour, Emerald, Bingley.

Vickery, B.C..C. & Vickery, A. (1992) Information science in theory and practice, Bowker-Saur.

Webster, F. (2014) Theories of the information society, 4th Ed., Routledge, London.

Weinberg, B.H. (1990) Regarding Michael Buckland's article *Journal of the American Society for Information Science* 41(6) 428-428.

Young, J.F. 1971, Information theory, Butterworths, London.

6.4 Appendix D: Reflections on the Dissertation

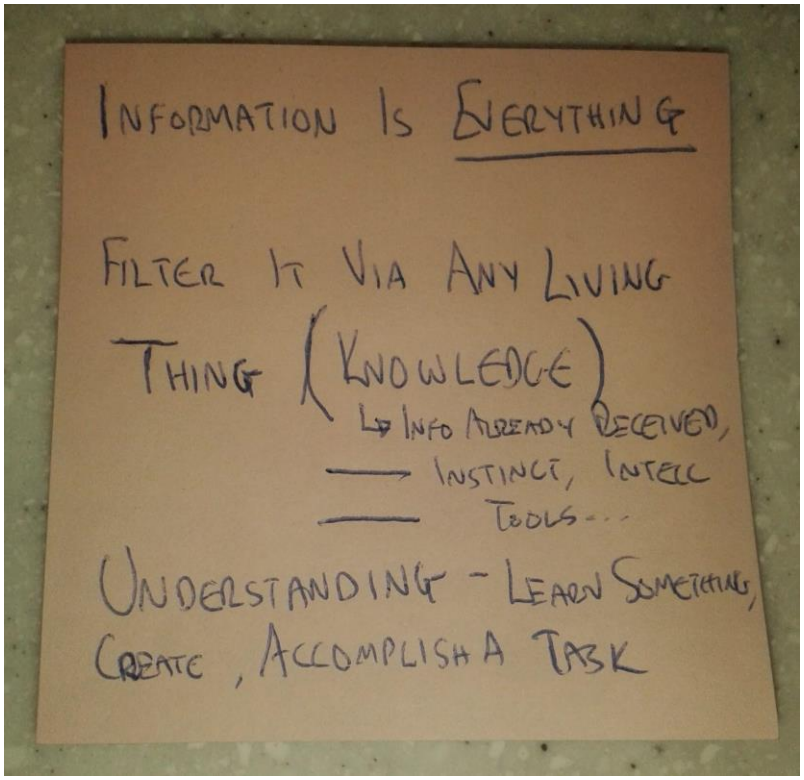

Figure 1: The Post-It Note that started it

At the start of the course this dissertation will bring to an end, the MSc Library at City University London, in an Introductory lecture we were addressed by a current PhD student who talked about his thesis. It was about information and how everything was information. Following this, and the first few lectures in a course called Library and Information Science Foundation, I sat down one evening and started to formulate my own theory of information. Figure 1 is the result.

If Information was everything, or in everything, it was surmised, then if that was filtered through knowledge you would gain understanding and learn or create something, perhaps, or accomplish a task.

After another few lectures, I found that my idea was hardly original and now I see that this is only a small part of the issue. However, I was hooked by information theory and kept looking for further confirmation of this idea in each module.

When colleagues asked me what I would do for my dissertation, I kept telling them, jokingly, that I would come up with my own theory of information. As it turned out, I would attempt something along those lines.

Overall, I have found the dissertation process very rewarding and interesting. However, I do wish that I had gone about it differently.

Taking the opportunity to float my idea in the final RECS (Research, Evaluation and Communication Skills) session would have given me some strong feedback and got me starting to think about what I was doing much earlier. As such, I wish that I had both done this and started my Dissertation Proposal much earlier, taking time to start the reading and develop the idea more.

Instead, it was rushed and I was left with a lot of catching-up to do once the proposal was agreed. This also meant that, upon conducting a better literature search, and from doing more research into the literature, I ended up not reading everything I listed in my proposal.

However, once I did get going, I found that each time I read some articles, the idea of how the dissertation would be formed grew more solid in my mind, which was a good feeling to get each time I worked. Having said that, I did not take advantage of having a supervisor enough. A danger of this dissertation was that I was very set in what I wanted to do and how I wanted to go about it from the start, and this didn't really change, but perhaps it would have if I had taken the time to discuss it with my supervisor or just asked questions as I went along.

In terms of the relation of the final results to the proposal - because I was so sure of my idea before starting, the dissertation did not alter too much. However, the more I read, the more I discovered was in the topic. I had thought about a division in LIS between computing and human approaches, but had not thought about the effect of truth.

I always felt that knowledge should be more overarching and appear in different places in the relationship of data, information and knowledge. However, I did not realise how many different ways there would be to express this, nor had I ever thought about emotion as playing a part - the article by Thellefsen et al (2014) had a particularly positive effect on me.

As I said, I always felt that knowledge should be more overarching and appear in different places but it was only late on in the process (after much of the dissertation had been written) that how this would look in a model really cemented itself in my mind. This I found quite surprising, really, but was pleased that it came then and was not something that I had decided on and sought to find in the literature - as it would have been if I had stuck with the original view in Figure 1.

Finally, I feel I have learned and gained many skills while writing this dissertation and will serve me well in my future career. They are: planning and organisation skills, information searching skills, writing and editing skills and project management skills.