

Spring 5-14-2015

# Continuing the DIKW Hierarchy Conversation

Judith Pinn Carlisle

*University of Michigan-Flint, carlislj@umflint.edu*

Follow this and additional works at: <http://aisel.aisnet.org/mwais2015>

---

## Recommended Citation

Carlisle, Judith Pinn, "Continuing the DIKW Hierarchy Conversation" (2015). *MWAIS 2015 Proceedings*. 8.  
<http://aisel.aisnet.org/mwais2015/8>

This material is brought to you by the Midwest (MWAIS) at AIS Electronic Library (AISeL). It has been accepted for inclusion in MWAIS 2015 Proceedings by an authorized administrator of AIS Electronic Library (AISeL). For more information, please contact [elibrary@aisnet.org](mailto:elibrary@aisnet.org).

# Midwest Association for Information Systems MWAIS 2015 Pittsburg, Kansas

## Continuing the DIKW Hierarchy Conversation

Judith Pinn Carlisle  
University of Michigan-Flint  
carlislj@umflint.edu

### ABSTRACT

This paper is written as a response to an invitation put forth by Jennex and Bartczak to continue the discussion of about the nature and role of the so-called *knowledge pyramid* given in the 2013 paper "A revised knowledge pyramid," (Jennex and Bartczak, 2013). In this paper, I briefly re-examine the origin of the data-information-knowledge-wisdom (DIKW) pyramid, refute two assertions made to support the Jennex-Bartczak revised pyramid, suggesting that the revised pyramid not be adopted as its own theoretical foundation is weak.

### Keywords

Knowledge Management, Knowledge Management System, DIKW Hierarchy, Knowledge Hierarchy

### INTRODUCTION

"What is knowledge?" is a question poorly dealt with in much of the knowledge management (KM) literature. In their influential paper, authors Alavi and Leidner, although they acknowledge that understanding various perspectives of knowledge may be useful, dismiss ramifications of philosophical approaches to understanding knowledge because lack of such understanding "... was neither a determinant factor in building the knowledge-based theory of the firm nor in triggering researcher and practitioner interest in managing organizational knowledge," (Alavi and Leidner 2001, p. 109). The position that focusing on definitions of knowledge is more problematic than valuable is a position clearly supported in some KM literature. For example (Keen and Tan, 2007) pursue a "corporatist" perspective of KM that is, that knowledge is viewed as an organizational asset write that

"There can never be a universal 'theory' of knowledge management, any more than there is any consensual agreement on what is knowledge in the mainstream of philosophy or any shared operational agreement as to its nature across the arts, sciences, theological, and political fields... if the conception of 'knowledge' remains a constant debating point and source of demurrals, no one gains neither KM pragmatists, philosophical idealists, nor activists in the ant corporatist sphere. The discussion just gets cloudier instead of clearer," (Keen and Tan, 2007, p. 2).

Still, unless knowledge management and knowledge management system researchers and developers continue to grapple with the hard questions of what knowledge is, it is difficult to put forth that a science underlies these systems.

The work presented in this paper reflects the following view. The creation and codification of any science is difficult, and sometimes seemingly impossible. However, this should not discourage researchers and developers but should serve as either a warning or an inspiration. The DIKW hierarchy is a model frequently used in the KM literature to explore the nature of knowledge and, based upon its frequent reference in the KM literature, serves as one theoretical foundation of KM. KM, as a subfield of information systems and an active and influential component of knowledge management system (KMS) research requires as sound a theoretical foundation as can be provided. In this paper, I take up the invitation given in (Jennex and Bartczak, 2013) to consider that proposed DIKW pyramid augmentation and explication as a robust model of KM. I identify and explore two shortcomings in

their proposed pyramid, ultimately recommending that this revised pyramid be excluded as a theoretical foundation of KM.

## REVIEW: THE KNOWLEDGE HIERARCHY

Sharma (2004) attributes the development of the knowledge hierarchy, at least within the context of knowledge management, to Ackoff (1989). The Ackoff paper itself does not give the source for its proposed hierarchical structure, presenting it as Ackoff's own ruminations and thoughts. I propose that Aristotle's Great Chain of Being is one of the first recorded depictions of hierarchical relationships between multiple diverse entities. The Great Chain of Being is divided into three levels. The third level depicts a hierarchy where each ascending level of the hierarchy, while retaining the characteristics of the level immediately subordinate to it, is superior to its subordinate. This is quite compatible with the hierarchical structure suggested by Ackoff, as he writes that "Wisdom is located at the top of a hierarchy of types," (Ackoff, 1989, p. 3) suggesting that as the highest level of the hierarchy of types, wisdom is somehow superior to the types below it.

Ackoff describes a hierarchical relation between five items, data, information, knowledge, understanding, and wisdom. Subsequent depictions of the knowledge hierarchy typically exclude the understanding level. Ackoff's hierarchy does not require that data transform into information, information into knowledge, or knowledge into wisdom. Instead, he states that each category is included in the next, "Each of these includes the categories that fall below it --- for example, there can be no wisdom without understanding and no understanding without knowledge," (Ackoff, 1989, p. 3). So, while Ackoff is typically identified as the source of the knowledge pyramid, he is not the source of the suggestion that through a series of transformations data becomes information which becomes knowledge which becomes wisdom.

Popular definitions of information, knowledge, and wisdom depict each as being composed of the type that resides on the pyramid directly below it with augmentation. That is, information is defined as data with additional characteristics, knowledge as information with additional characteristics, and wisdom and augmented information. Typical of much of the literature, Jennex and Bartczak characterize data as "... Basic, discrete, objective facts such as who, what, when, where, about something," (Jennex and Bartczak, 2013, p. 20). The definition provided for information adds the attribute of context to the definition of data. Knowledge is defined as information (contextualized data) understood within a culture emphasizing what they identify as "the why about something," (Jennex and Bartczak, 2013, p. 20) or that provides insight and understanding. In other words, knowledge is information about the some thing's cause or purpose or other insight into the thing at under consideration. Wisdom is this culturally understood and motivated knowledge placed into an additional, but unidentified framework. The authors add that wisdom may be knowledge within the context of a nomological network allowing, per them, non-intuitive correlations between various pieces of knowledge to emerge.

## THE JENNEX-BARTCZAK REVISED PYRAMID

Jennex and Bartczak (2013) utilizes definitions of data, information, knowledge, and wisdom which reflect a hierarchical relationship between the four items. They define data as a collection of facts. Information, defined in terms of its relationship to data, is composed of data that are related to one another via a context of some sort. Knowledge, is defined as culturally understood information. Wisdom is knowledge placed within a framework or a nomological network. It is not uncommon for KM authors, for example (Broadbent, 1998; Martin, 2008; Meadow and Yuan, 1997; Morrow, 2001) to define the four types as they relate to one another. This approach supports the following beliefs

- some essential relationship holds between the four entities
- each type transforms to a different type, that is
  - + information is made of data,
  - + knowledge is made of information, and
  - + wisdom is made of knowledge.

This paper specifically addresses two supporting conditions of the revised DIKW pyramid proposed by Jennex and Bartczak. Each of these conditions weakens the theoretical foundation of the revised pyramid. Counter-examples are given for each condition.

One argument Jennex and Bartczak (2013) uses to motivate the revised pyramid has to do with the amount of information that can be derived from a data set. Jennex and Bartczak (2013) argues that the amount of information emerging from a data set  $x$  is at least  $x!$ . The paper states "If information is the structuring of data into meaningful combinations, then the number of possible combinations for a quantity  $x$  of data is minimally  $x!$  implying that there is possibly a greater amount of information than the original amount of data," (Jennex and Bartczak, 2013, p. 22). A certain ambiguity exists with the written statement. The *if-then* construction of the sentence implies that the two phrases *meaningful combinations* and *possible combinations* are somehow related. The idea that various arrangements of data can result in information harkens to such basic definitions of information given in typical information systems textbooks. For example, "By information we mean data that have been shaped into a form that is meaningful and useful to human beings," (Laudon and Laudon, 2014, p. 15). It appears that Jennex and Bartczak (2013) suggests that because the number of permutations arising from a data set with a cardinality of  $x$  is  $x!$ , then the number of some significant portion of those permutations will be greater than  $x$ .

This assertion is true for certain conditions of permutation. Conditions can exist where such a relationship between a data set and the resulting information derived from that data set can be described as a factorial function over the cardinality of the data set. However it is not necessarily true that the number of all information items derived from a data set generated by  $x!$  will be greater than the cardinality of  $x$ . The most obvious case for this to be true is the case when every permutation corresponds to a well-formed representation.

I offer two different examples, one to illustrate the case where more information does arise from the manipulated data set than the cardinality of the foundational data set and second given as a counter-example to illustrate the case where the emerging information items number less than the original data set. I suggest that this counter-example weakens the assertion that as movement is made up the pyramid from data to information to knowledge to wisdom the amount of membership for each type increases.

My first example uses a weighted position code approach to depicting numerical values. Weighted position code approaches to numerical representation offer a system where all sequences of numerals reference a specific value and so can be said to have a meaning. A weighted position code approach to value representation can be described as:

Where the base is given as the radix ---  
 $V$  = the depicted value  
 $b_i$  = the value of a digit in position  $i$  is given  
 $n$  is the number of digits to the left of the radix point  
 $m$  = the number of digits to the right of the radix point

Through such an equation I can depict the value of the binary number 101.1 as

$$\begin{array}{r} (1 * 2^2) + (0 * 2^1) + (1 * 2^0) + (1 * 2^{-1}) \\ (1 * 4) + (0 * 2) + (1 * 1) + (1 * 1/2) \\ 4 + 0 + 1 + 1/2 \\ 5 \frac{1}{2} \end{array}$$

Hence given a set  $p$  of three numerals  $p = \{1, 2, 3\}$ , then all six permutations reference a value, hence are meaningful as all numerals have meaning in all positions:

123	213
132	312
231	321

Yet consider that, given a different type of formal system using arbitrary symbols to encode information, the same relationship does not hold. Consider the set  $p'$  where  $p' = \{a, b, c\}$ . Six permutations emerge:

abc	bca
acb	cab
bac	cba

However, adding a context, such as "the English language" results in an information set whose members are those permutations that are meaningful, that is combinations of letter tokens that create real English words. Only two of these permutations have the potential to be meaningful, *bac* and *cab*. Of these two possibilities, only one, *cab*, is a recognized English word and so, meaningful. While it is possible that other permutations may be words, and so meaningful, in other languages, in the context of the English language I see that the cardinality of the information set that emerges from the data set  $p'$  is less than the cardinality of the set. That is, the factorial function does not describe the relationship between the Roman alphabet and the information sets that emerge from random collections of letters. This holds for the English language. The Roman alphabet has 26 characters.  $26!$  is  $4.032914611266057 \times 10^{26}$ . The number of meaningful permutations is much lower. The current version of the Oxford English Dictionary has approximately 600,000 words<sup>1</sup> --- far below the possible permutations.

As the information context becomes more constrained, fewer and fewer information candidates emerge from a given data set. Consider that if the information context is altered from identifiable words to syntactically well-formed inscriptions, the resulting information set is quite sparse, moving from syntactically well-formed inscriptions to semantically well-formed inscriptions the information set becomes even smaller. For example, given the data set  $V = \{\text{some, girl, flies}\}$  and a generative syntax expressing the English language, the following combinations emerge:

some girl flies	some flies girl	girl some flies
girl flies some	flies some girl	flies girl some

Given the constraints of the formal language system, only one syntactically well-formed candidate emerges: *some girl flies* and no semantically well-formed candidate emerges.

We see that the the amount of information that emerges from a data set is affected by how constraining the context condition is. As the information context becomes more constrained, the amount of information emerging from a data set appears to decrease. In the examples above, the weighted position code approach to value representation has few constraints. In any base, any valid numeric symbol can reside in any position without violation. With such low constraints, the cardinality of the information set  $y$  emerging from data set  $x$  can be  $x!$ . This is not true when the constraint conditions become more confining as in the case of language. Mitigating rules arbitrate which position any letter token or word token, with relation to the other letters or words with which it co-occurs, can occupy. The cardinality of the information set  $p$  emerging from data set  $q$  will be less than  $q!$ .

The second aspect of information offered by Jennex and Bartczak repeats an idea given in the Houston and Harmon paper "Re-envisioning the information concept: systematic definitions," (Houston and Harmon, 2002). The authors Jennex and Bartczak somewhat alter Houston and Harmon's depiction of the relationship between data and information, distilling it to the sequence of equations:

$$I = \Sigma(D), K = \Sigma(I) = \Sigma\Sigma(D), \text{ and } W = \Sigma(K) = \Sigma\Sigma(I) = \Sigma\Sigma\Sigma(D)$$

That is, information is the result of some summation function over data, knowledge of the same summation function over information, and wisdom as summation over knowledge. There are some distinct limitations of these depictions as all summations are absent an index of summation, and range over which the summation proceeds. As stated (Jennex and Bartczak, 2013) appears to vary from Houston and Harmon (2002)'s explanation which eliminates a temporal dependence from data yet yields such a dependence for information. To illustrate the weakness of Jennex and Bartczak's depiction, I also alter Houston and Harmon's example. However, I explicitly state that I acknowledge a temporality as an aspect of data but deny such a temporal aspect for information.

<sup>1</sup> (given at the Oxford English Dictionary website <http://public.oed.com/about/>)

Data exist in time and space. That data exist at some time and do not exist at others is a necessary component of the ability to observe the presence of data as well as its absence. It is also a necessary component marking the ability to cause data to occur, alter specific characteristics of various data at specific points in time, as well as the ability to destroy data. However, information has no such temporal component. Using ordinary language, I can show that is nonsensical to assert things such as:

- The information will be over in an hour.
- I'll see you after the information.
- I'll meet you near the information.

Given the idea that information consists of augmented data, then information should possess all of the characteristics of its contributing data, plus additional characteristics. However, since information fails to include the characteristics of existing in time and space, stating that information emerges as a summation of all contributing data entities seems to be a misnomer.

## CONCLUSION

Jennex and Bartczak (2013) questions the existing DIKW hierarchy and propose a modification intended to create a more robust and descriptive DIKW hierarchy. In this paper I have shown that two assumptions underlying the revised DIKW hierarchy are inconsistent and do not necessarily provide the theoretical basis required of them by (Jennex and Bartczak, 2013) calling the validity of the revised pyramid which rests upon them itself into question.

## REFERENCES

1. Ackoff, R.L. (1989) From Data to Wisdom Presidential address to ISGSR, June 1988, *Journal of applied systems analysis*, 16, 3 - 9.
2. Alavi, M. and Leidner, D. (2001) Review: knowledge management and knowledge management systems: conceptual foundations and research issues, *MIS Quarterly*, 25, 1, 107 - 136.
3. Broadbent, M. 1998 The phenomenon of knowledge management What does it mean to the information profession? *Information Outlook*, 2, 5, 23 -36.
4. Houston, R.D. and Harmon, E.G. (2002) Re-envisioning the information concept: systematic definitions, in H. Bruce, R. Fiedel, P. Ingersen, and P. Vakkari (Eds), *Emerging Frameworks and Methods, Proceedings of the fourth International Conference on Conceptions of Library and Information Science (CoLIS4)*, July 21 - 25, Seattle, WA, USA, 305 - 308.
5. Jennex, M.E. and Bartczak, S.E. (2013) A revised knowledge pyramid, *International Journal of Knowledge Management*, 9, 3, 19 - 30.
6. Keen, P.G.W. and Tan, M. (2007) Knowledge Fusion: A framework for extending the rigor and relevance of knowledge management, *International Journal of Knowledge Management*, 3, 4, 1 - 17.
7. Laudon, K.C. and Laudon, J.P. (2014) *Management Information Systems: managing the digital firm*, 13th edition. Pearson Publishers, Upper Saddle Valley, NY.
8. Martin, B. 2008. Knowledge management, in C. Blaise (Ed.), *Annual review of information science and technology (ARIST)*, vol. 42, 371-424, Information Today, Medford, NJ.
9. Meadow, C.T. and Yuan, W. 1997. Measuring the impact of information, *Information Processing and Management*, 33, 6, 697 - 714.
10. Sharma, N. (2004). The origin of the "data information knowledge hierarchy. URL typically given as [http://www.personal.si.umich.edu/~nsharma/diwk\\_origin.htm](http://www.personal.si.umich.edu/~nsharma/diwk_origin.htm), with retrievals at June 17, 2008, but no longer available at this site.

