Transactions of the SDPS:
Journal of Integrated Design and Process Science
21 (4), 2017, 1-4
DOI 10.3233/jid-2018-0006
http://www.sdpsnet.org

EDITORIAL

Building Intelligence in Digital Transformation

Jingwei Huang [*]

Department of Engineering Management and Systems Engineering, Old Dominion University, USA

We are in an age of great changes! An on-going profound societal movement is digital transformation, which is the pervasive diffusion of digital technologies in business and other societal processes. On the one hand, digital transformation is driven by the confluence of technologies such as Internet of Things, 3D printing, Big Data, machine learning and artificial intelligence (AI); on the other hand, digital transformation provides fertile fields for the further development of those technologies, and most importantly, for the growth of new technologies and new digital business, such as digital twins, blockchain, cryptocurrencies, digital archives, and smart contracts. Related to the emerging digital transformation phenomena, we expect many opportunities, challenges, and uncertainties. One thing for sure is that digital transformation will continue to make Big Data bigger and bigger, and in the emerging digital society, successes will belong to the organizations who are able to build intelligence in digital transformation by leveraging AI to turn data into knowledge and to use the knowledge in actions.

What are approaches to building intelligence in digital transformation? First of all, what are data? What are Big Data? What is knowledge? What are the differences between them? What do Big Data mean to us when we try to understand and model a complex system or a complex process? Do Big Data carry all information about a system or a process? The study of the nature of knowledge can be traced back to Plato's Theaetetus. In Philosophy, an influential definition of knowledge is "justified true belief" (Stanford, 2017). In AI, based on Kripke possible world semantics, a series of axiomatic systems (T, S4, S5, and KD45) to define knowledge and belief have been studied (Halpern and Moses, 1992). In information science and management science, many researchers have explored the concepts of data, information, knowledge, and wisdom, and their hierarchical structure - DIKW pyramid (Ackoff, 1989; Rowley, 2007). Bellinger et al (2004) illustrate the relations among data, information, knowledge, and wisdom in a two-dimension space (connectedness, understanding), where from data towards the direction of wisdom, the levels of understanding and connectedness increase. Intelligence is the ability to transform data into information and knowledge, to lift knowledge into wisdom, and to use knowledge in problem solving.

Big Data has been characterized with 3Vs – high-volume, high-velocity and high-variety by Gartner (2012). Volume is about the size or scale of data. Velocity is about the high speed in which data are produced and need to be processed. A common understanding of "variety" is about the various forms of data, such as structured and unstructured, text, images, videos, sensor readings, and so on; however, we must consider variety in another dimension – many different sources/views of data. The data about the same

[*] Corresponding author. Email: j2huang@odu.edu

 Published by IOS Press

object could come from different observers' radar, sonar, lidar, and so on. The data about the same individual could come from different social media. The new reports about the same event could come from different witnesses, journalists, and media agencies. This new dimension of "variety" brings in complexity beyond technologies. Relevant to this, IBM introduced the fourth V – "veracity" to address the uncertainty, reliability, and trustworthiness of data (IBM, 2013). Researchers have been adding in new Vs, such as "value" (high value of collective data and low value density of data). Back to the earlier questions, in the context of Big Data, *data can be interpreted as observations by a viewer from a specific view at a specific point of time*, thus representing an agent's belief. Knowledge is justified true belief in the relations, features, properties, patterns, and laws about the studied system or process, obtained through learning and logical reasoning including deductive, inductive, abductive, and analogical reasoning, based on known knowledge and the given trusted observations (data). It is important to evaluate the trustworthiness of those observations from different sources based on their provenance (Huang, 2008). People may remember the cheer "let's put everything in and let the data speak for itself", after some early successes of Big Data analytics, e.g. Google Flu Trend revealed influenza cases spiking two weeks prior to the CDC report in early 2010. As addressed earlier, data are observations by a viewer from a specific view at a specific point of time; Big Data provide much more observations than before; however, usually Big Data do not carry all information about a complex system or process. In a system's state space, Big Data may have high redundancy in some regular system states, and may have just few pieces of information (even no at all) in those rare system states. In the latter case, Big Data carry just little information. It is fair to say that Big Data technology provides us the opportunities to gather unprecedented scale and unprecedented finer time intervals of observations, thus offering larger opportunities to capture the truth about the studied system or process, e.g. the most recent finding of Higgs boson decay (Nature News, 08/31/2018) and the earlier discovery of Higgs at LHC (Nature News, 07/04/2012), in which a scientist explained "Out of some 500 trillion collisions, 'the signal we're looking at is some tens or dozens of particles', ... The feat is equivalent to picking a few grains out of an Olympic-sized swimming pool full of sand." Surely, the increased density of experimental data collection powered by improved computing power directly contributed to the existing discovery.

Jim Grey had a vision that science has had three research paradigms in history, that is, empirical (characterized as based on observation and description of natural phenomena), theoretical (characterized as based on mathematical models), and computational (characterized as based on simulations, when theoretical models become too complex to derive and prove propositions); now, today's science is transforming into its forth paradigm – data-intensive paradigm, which leverages Big Data and associated Big Data technologies in scientific discovery. This insight has inspired various data-driven research in many scientific and engineering disciplines, including systems design and process science. For example, the recent DSR 2018–Workshop on Data Driven Design and Learning, held at Montreal, Canada. Data-intensive science and engineering is a manifestation of digital transformation in the fields of science and engineering.

Simon (1997) revealed that human cognitive activities have "bounded rationality", i.e. the "rational choice that takes into account the cognitive limitations of the decision maker - limitations of both knowledge and computational capacity". As we see from the earlier stories, the increased computing capacity in the past has significantly extended human society's capability in knowledge acquisition and knowledge-based decision making. The first wave of lift in computing power was mainly attributed to the hardware progress dominated by Moore's Law (Moore, 1965) for almost 50 years, which is now close to saturation; the second wave, now right in its peak, is from parallel and distributed computing, represented by Cloud Computing; the third wave, which now is coming, is probably from artificial intelligence, to make computing more intelligent and more efficient.

Now let us return to our first question regarding how to build intelligence in digital transformation. Digital transformation includes digitization and digitalization. Digitization is to create a digital representation of a physical or analog product; examples include the digital version of a paper book, a digital description of a product sold at Amazon, and a part design produced from CAD. An advanced example reflecting the state-of-the-art development could be digital twins of physical machines or products, which have been used in digital manufacturing. Digitalization is about to use digital technologies to operate digital or digitized entities in societal processes including business processes. Examples include the digital business of Netflix and Amazon, digital supply chain, and others. The Digital Manufacturing and Design Innovation Institute (DMDII) of UI LABS at Chicago, USA is an excellent showcase of digital transformation in manufacturing industry. Digitalization will lead to digital business. Digital entities and digitized entities provide the space to employ various digital technologies and to leverage Big Data and AI technologies. Currently, digital transformation in business has been largely attracted to Big Data analytics and machine learning. We believe, beyond machine learning, AI has immense potential to improve digital business in many aspects, naming a few, intelligent designers, intelligent business agents, intelligent process monitors, intelligent resource managers, and so on.

This issue of JIDPS presents five research papers. Each of them reflects some interesting aspects of the related research issues. In digital transformation, a significant portion of data are in the form of unstructured text. Therefore, efficient knowledge discovery from text and information retrieval from text are critical to digital transformation. Paper “A Semantically Enriched Context-Aware Stemming Algorithm”, by Melyara Mezzi, Nadjia Benblidia and Xiangji Huang, proposes a hybrid stemming algorithm to combine the strength from rule-based stemmer and dictionary stemmers. Stemming is an essential pre-processing stage for text mining; stemming transforms a word from its variant forms into its root form (or word stem). The authors show the improvement of their algorithm by comparing with some representative algorithms with benchmark datasets.

Big Data may come from many different sources even just within an enterprise. Echoing this feature of “variety”, paper “Knowledge Management Through Ontology-Driven Integration of Disparate Knowledge Sources”, by Jan Werrmann and Bernd J. Kramer, addresses the issues on enterprise knowledge search from disparate sources. Knowledge and information relevant to a specific case or situation are distributed in many different information systems, knowledge bases, databases, wikis, emails, and document archives, and it is a challenge to find the knowledge related to a given scenario at real time. The authors propose to represent knowledge/information items and their relations as ontologies, and to construct an ontology directed quick search. They also allow dynamic reinforcement of the relations among those knowledge/information items by using users’ case specific searches and feedbacks. They also report their successful test of the prototype implementation on the business of car maintenance, repair, and services by a car manufacturer.

One of the benefits of digital transformation is the flexibility of digital products and services, which allows to offer personalized services. Paper “Integration of Digital Social Story Intervention into Differentiated Instruction Framework”, by Win Ko Min and Lau Bee Theng, illustrates how they effectively leverage digital materials augmented with personalized instructions in special education for children diagnosed with Autism Spectrum Disorder.

To build intelligence in digital transformation, not only the application of machine learning and AI is important, but also the education, knowledge and skills preparation, and fostering culture for digital transformation are equally important. Paper “Empowering Rural Youth for Socio-Economic Benefits: A Case Study of Knowledge Management Practices in Sarawak”, by Wan-Tze Vong, Patrick H. H. Then and Tien-Hiong Teo, introduces the design of the Rural ICT-Guided Home-based Technopreneur (RIGHT) program for training rural youth in Malaysia on information and communication technologies, from knowledge management perspective with components of knowledge acquisition, knowledge utilization, and knowledge sharing; and the authors evaluate the effectiveness of this special knowledge management process.

Different from many ways of knowledge acquisition by learning and reasoning, paper “Application of Design Methodologies to Web System Design: A Case Study of JIDPS Editorial System”, by Mengli Shu, Suo Tan, Liang Fu, Yijing Zeng, Xinlin Cao, and Yong Zeng, shows in the context of product design an interesting approach to acquire knowledge (specifically, customer requirements) from a conversation dominated by a general structure called Recursive Object Model (ROM). The authors integrate this intelligent requirement acquisition process into a "Quality Function Deployment" based product design framework. The authors demonstrate how this approach works with a case study on the design of a web-based journal editorial system.

In the on-going digital transformation, several disruptive technologies such as Internet of Things and blockchain are still in their infant stages. Internet of Things extends the Internet into the physical world. We all already experienced the remarkable changes that Internet brings to us; but the impact from Internet of Things will be much stronger. The applications of Internet of Things in industry are leading to the "fourth industrial revolution" (Schwab, 2017). This new development is changing the landscape of engineering systems design (Huang, 2016). The dramatic and active interactions between digital technology development and the digital transformation in society will continue. JIDPS will continue to publish articles on the theme and relevant emerging research topics.

References

Ackoff RL. (1989). From data to wisdom. Journal of applied systems analysis. 1989 Jun; 16(1), 3-9.

Bellinger, G, Castro, D and Mills, A (2004). Data, Information, Knowledge, and Wisdom. http://www.systems-thinking.org/dikw/dikw.htm

Gartner (2012), Big Data, https://www.gartner.com/it-glossary/big-data

Halpern JY, Moses Y (1992). A guide to completeness and complexity for modal logics of knowledge and belief. Artificial intelligence. 1992 Apr 1; 54(3), 319-379.

Huang, J. (2008). Knowledge provenance: An approach to modeling and maintaining the evolution and validity of knowledge (Doctoral dissertation), University of Toronto.

Huang, J., Seck, M. D., & Gheorghe, A. (2016). Towards trustworthy smart cyber-physical-social systems in the era of internet of things. In IEEE SOSE 2016.

IBM, (2013). Analytics: The real-world use of big data, https://www-935.ibm.com/services/multimedia/Analytics_The_real_world_use_of_big_data_in_Financial_services_Mai_2013.pdf

Moore, Gordon (1965), Cramming More Components Onto Integrated Circuits. Electronics. Volume 38, Number 8, 114-117.

Nature News (07/04/2012). Physicists declare victory in Higgs hunt.

https://www.nature.com/news/physicists-declare-victory-in-higgs-hunt-1.10940

Nature News (08/312018). LHC physicists finally uncover Higgs 'bottom' decay. https://www.nature.com/articles/d41586-018-06130-9?utm_source=briefing-dy&utm_medium=email&utm_campaign=briefing&utm_content=20180831

NIST (2015). NIST Big Data Interoperability Framework: Volume 1, Definitions (No. Special Publication (NIST SP)-1500-1).

Rowley J. (2007). The wisdom hierarchy: representations of the DIKW hierarchy. Journal of information science. 2007 Apr; 33(2), 163-80.

Simon, H. A. (1997). Models of Bounded Rationality. Vol. 3. The MIT Press.

Stanford Encyclopedia of Philosophy (2017). The Analysis of Knowledge. https://plato.stanford.edu/entries/knowledge-analysis/

Schwab, K. (2017). The fourth industrial revolution. Crown Business.