AI ethics and its impact on knowledge management

Anthony J. Rhem¹

Received: 12 September 2020 / Accepted: 17 September 2020 / Published online: 6 October 2020
© Springer Nature Switzerland AG 2020, corrected publication 2020

1 Introduction

The focus of knowledge management (KM) is to enable people and organizations to collaborate, share, create and use knowledge. Understanding this KM is leveraged to improve performance, increase innovation and grow the knowledge base of both people and the organization. Knowledge must be Dynamic, Accurate and Personal to be applied in the decision-making process. Artificial Intelligence (AI) through machine learning allows machines to acquire, process and use knowledge to perform tasks and to unlock knowledge that can be delivered to people to improve the decision-making process. AI plays an important part to delivering knowledge in a digitized organization by elevating how the delivery of knowledge occurs to the people who need it. AI is used to scale the volume and effectiveness of knowledge distribution. It is imperative that when AI is applied to deliver knowledge for people to make decisions; including when AI is used to make decisions without human involvement; that the knowledge is without bias and the decisions made with the knowledge are ethical.

1.1 Incorporating artificial intelligence (AI)

AI provides the mechanisms to enable machines to learn. Incorporating AI in the delivery of knowledge will facilitate fast, efficient and accurate decision making. AI provides the capabilities to expand, use, and create knowledge in ways we have not yet imagined. AI systems which use machine learning, can detect patterns in enormous volumes of data and model complex, interdependent systems to generate outcomes that improve the efficiency of decision making. The use of AI (machine learning) in delivering knowledge is based on the data that is used to train the machine learning

algorithms. We must keep in mind that when it comes to AI, we need both responsible use and responsible design.

2 Delivery of knowledge

Knowledge Management (KM) in organizations is based on an understanding of knowledge creation, knowledge transfer, knowledge use, knowledge flows, and knowledge governance. In implementation, KM is an effort to benefit from the knowledge that resides in an organization using it to achieve the organization's mission and the activities of its users. The transfer of tacit or implicit knowledge to explicit and accessible formats, the connection of tacit to explicit knowledge (which is connecting your organization's experts to explicit knowledge), and the understanding of social network analysis within the organization contributes to the effective distribution and use of knowledge. It is important to not only understand what knowledge management means for the organization, but also understand the nature of the organization's knowledge itself.

AI effect on KM is to provide the delivery of knowledge in a Dynamic, Accurate and Personal way. Dynamic knowledge is constantly updated adhering to your organization's content (information and knowledge) lifecycle management processes. This also includes the experts who can provide insights about the knowledge. The Dynamic component of knowledge reflects your organization's brand, tone and evolves over time. The Accurate component of knowledge is identified as the authoritative source and authoritative voice for that subject matter. This knowledge is accepted by your organization as the source of truth. The Personalized component of knowledge answers the questions that the users of your knowledge are seeking. Personalized knowledge is tailored to what the individual needs to make a decision and presented in a way that is tailored to the device and applications the users are leveraging to access answers to their questions. Personalized knowledge is facilitated by how knowledge flows throughout your organization.

✉ Anthony J. Rhem
tonyr@ajrh.com

¹ A. J. Rhem and Associates, Inc., 205 North Michigan Ave., Suite 810, Chicago, IL 60601, USA

Using AI in the delivery of knowledge, let's examine the characteristics of Dynamic, Accurate and Personal knowledge delivery:

2.1 Dynamic

Knowledge is a result of a varied set of processes and flows, which demonstrate the active nature of knowledge. The dynamic nature of knowledge stems from the fact that knowledge is active and always changing. This change is a direct result of the changing and evolving human experience at your organization. This also effects the relationships (connections) between employees and their experience that is gained through assignments, training and learning.

To manage the dynamic nature of knowledge, it must be governed and maintained. Knowledge must be constantly updated adhering to your organization's content (information and knowledge) lifecycle management processes. This also includes the experts who can provide insights about the knowledge. It must allow for the evolution of changing experiences and connections through communities and mapping the connections of people to people, people to content and content to content. The Dynamic component of knowledge reflects your organization's brand, tone and evolves over time.

2.2 Accurate

The Accurate component of knowledge is identified as the authoritative source and authoritative voice for that subject matter. This knowledge is accepted by your organization as the "source of truth". The level of accuracy of knowledge determines the quality of performance of the AI framework in real-world situations. The use of standards in the creation, use and maintenance of knowledge will serve as a basis for consistent and accurate knowledge. The importance of standards will provide people and organizations with a basis for mutual understanding and are used as tools to facilitate communication and measurement of the quality and accuracy of the knowledge within the organization.

AI will be utilized to allow users to collaborate, communicate and distribute all company knowledge and information internally (and through external channels). AI provides in-depth, instant, on-demand knowledge to your users. In doing so, the knowledge, including all connections and content (text, images, voice, video, etc.) must be kept up-to-date and accurate. The knowledge that is accessed must be the "source of truth" and trusted by all users.

2.3 Personal

The Personalized component of knowledge answers the questions that the users of your knowledge are seeking.

Personalized knowledge is tailored to what the individual needs to make decisions and presented in a way that is tailored to the device and applications the users are leveraging to access answers to their questions. Personalized knowledge is facilitated by how knowledge flows throughout your organization. With AI, knowledge is optimized for the benefit of the individual user. AI will facilitate the mapping of knowledge (tacit and explicit) to enable everyone seeking knowledge within the organization to be provided its necessary context.

As knowledge within the organization grows, fast response and personalized access to organizational knowledge assets have become necessary to execute and deliver results. This is particularly important for content-related products and services, such as consulting, marketing/advertising, and communications within an organization. The digital nature of these products allows for more customized delivery within the AI Framework. To provide personalized services a complete understanding of user profiles and accurate associations to people and are essential.

3 Scale the delivery of knowledge

AI plays an important part in delivering knowledge in a digitized organization by elevating how the delivery of knowledge occurs to the people who need it. AI is used to scale the volume and effectiveness of knowledge distribution by:

- Predict trending knowledge areas/topics that your employees need
- Identify which targeted knowledge will resonate with your employees based on real-time engagement and content consumption
- Auto-curate and personalize knowledge based on individual preferences
- Improve content decisions by leveraging machine learning around what content will be best suited to address the situation
- AI will make search and its search products more relevant, precise and efficient.
- AI through intents will be able to better know what content your employees need. Intents will provide a better understanding of what the employee is looking for by better understanding the intended use of the content.
- Chatbots with Natural Language Processing (NLP): will provide value for all employees in the various organizational functions along critical decision-making points with personalization of the delivery of knowledge. Chat Bots with NLP will provide cognitive capabilities to understand, interpret and manipulate human language that will enable the bots to anticipate the needs, attitudes and aspirations of users to aid in decision making and

improve outcomes, all geared to achieve substantial business value.

4 Ethical issues of AI delivery of knowledge

There are various ethical issues that arise with certain uses of AI technologies. Each type of AI technology will raise different types of ethical issues when it comes to making decisions based on AI. AI technologies include, text analysis, natural language processing (NLP), logical reasoning, game-playing, decision support systems, data analytics, predictive analytics, autonomous vehicles, and Digital Assistants (chatbots), just to name a few. AI also involves a number of computational techniques such as classical symbol-manipulating inspired by natural cognition, or machine learning via neural networks.

Some examples of the decisions being made with knowledge that is provided by AI applications include:

4.1 Autonomous vehicles

Autonomous vehicles (AV) manufacturers have collected vast amounts of data over the course of training their AV algorithms while operating their AV prototypes. The real-time driving data that AV developers collect is proprietary and not shared across firms. This makes the data difficult to share and the accessibility of this data challenging. The ethical and sensitive nature of this data has to be protected. This data can in many cases include location and user behavior information that is used for the vehicles to make driving decisions. This data will also need to be managed and protected. In addition, the shift to AVs could also have a significant effect on freight, taxi, delivery and other service jobs. A rapid transition to AVs in the industry from a profit-maximizing perspective can be expected especially as the technology advances. This technological shift will displace workers, highlighting the need for policy work focused on skills and jobs in the context of a transitioning work environment [1, 2].

4.2 Financial services

In financial services (including insurance companies) many organizations are deploying AI solutions. In many cases financial service organizations are combining different AI solutions combined with machine learning (i.e., Robotic Process Automation, Language processing/NLP, Deep Learning Decision solutions) to deliver knowledge to its users to make better decisions. There are several benefits to deploying AI solutions in the financial services sector, which include improving customer service, providing smarter investment tools, credit analysis and scoring, and smarter

financial analysis tools. However, these AI empowered tools raise policy questions related to ensuring accuracy, preventing discrimination and bias (especially in credit analysis and scoring) as well as impacting jobs.

As it pertains to AI being used for credit analysis and scoring, the difficulty to explain results from these algorithms have been a problem (OECD Artificial Intelligence in Society, 2019). This is driven by the legal standards in several countries, including the United States that require high levels of transparency. For example, in the United States the Fair Credit Reporting Act (1970), and the Equal Credit Opportunity Act (1974) implies that the process and the output of any algorithm has to be explainable.

4.3 Healthcare

AI applications in healthcare and pharmaceuticals has produced many benefits by delivering knowledge to detect health conditions early, deliver preventative services, optimizing clinical decision making, discovering new treatments and medications, delivering personalized healthcare, while providing powerful self-monitoring tools, applications and trackers. Although AI in healthcare offers many benefits it also raises policy questions and concerns that include access to (health) data and privacy, which includes personal data protection.

The healthcare sector is a knowledge-intensive industry and it depends on data and analytics to improve the delivery of healthcare (treatments, practices, procedures). There has been tremendous growth in the range of information collected, including clinical, genetic, behavioral and environmental data, with healthcare professionals, biomedical researchers and patients producing vast amounts of data from an array of devices (i.e., electronic health records (EHRs), genome sequencing machines, high-resolution medical imaging, and smartphone applications). How this data is collected used and protected will also bring challenges that all countries driven by its legal standards will have to address. In the United States organizations such as the Food and Drug Administration (FDA) and policies such as the Health Insurance Portability and Accountability Act (HIPAA) of 1996 are in place to ensure standards, guidelines, data security and privacy is adhered to and enforced.

5 Ethicality of AI applications

The ethicality of AI applications must be examined to understand if the outcomes of that specific application of AI is fully understood and it's not violating our (human) moral compass. The most immediate concern for many is that AI-enabled systems will replace workers across a wide range

of industries. AI brings mixed emotions and opinions when referenced in the context of jobs.

However, it's becoming increasingly clear that although AI may replace some jobs, it will create others. This will create the need to re-skill the workforce to fill these new jobs being created. Research and experience are showing that it's inevitable that AI will replace entire categories of work, especially in transportation (through autonomous vehicles), retail, government, professional services employment, and customer service. Conversely, companies will have the workforce that will be re-skilled to take on much better, higher value tasks.

The International Standard for AI delivered by the Organization for Economic Co-operation and Development (OECD), lays out general tenants for AI implementation that is focused on ethical adherence and the adherence to the well-being of humanity. There were 50 members comprising of governments, academia, business, civil society, international bodies, the tech community and trade unions that contributed to the standard [1, 2]. The following OECD AI value-based principles and recommendations for public policy and international cooperation (see below) are intended to be guiding principles for governments, organizations and individuals in the design, development and implementation of AI systems. The purpose for these guiding principles is to ensure that the public and users of AI best interest are primary and that the individuals designing, developing and implementing AI systems are held accountable to the OECD AI standard. According to the OECD Secretary-General Angel Gurría, these principles “will be a global reference point for trustworthy AI so that we can harness its opportunities in a way that delivers the best outcomes for all” [1, 2].

5.1 OECD AI value-based principles

The following is a synopsis of the OECD AI value-based principles.

- AI should benefit people and the planet by driving inclusive growth, sustainable development and well-being.
- AI systems should be designed in a way that respects the rule of law, human rights, democratic values and diversity, while applying the appropriate safeguards to ensure a fair and just society.
- There should be transparency and responsible disclosure around AI systems especially when they are engaging with the public.
- AI systems must function in a robust, secure and safe way throughout their lifetimes, while accessing and managing potential risks.
- Organizations and individuals developing, deploying or operating AI systems should be held accountable for their

proper functioning in line with the AI value-based principles.

6 The need for a people-centered AI approach

People-centered AI focuses on improving the human condition. The central theme of the OECD AI standard is to support, guide and influence policies that will enable AI applications to be people-centered. People-centered AI applications support the inclusivity and well-being of people that it serves; respects human-centered values and fairness; AI applications must be designed, developed and implemented with transparency; AI applications are robust and safe; and there is accountability for the results and decisions that AI applications produce and/or the decisions AI applications influence [1, 2].

6.1 Inclusivity and well-being

AI development must be done to ensure inclusivity and well-being. Artificial intelligence (AI) plays an increasingly influential role. As the technology diffuses, the potential impacts of its predictions, recommendations or decisions on people's lives increase as well. The technical, business and policy communities are actively exploring how best to make AI people-centered and trustworthy, maximize benefits, minimize risks and promote social acceptance.

AI has the potential to exacerbate the inequality and divides that exist concerning AI resources, technology, talent, data and computing power. In addition, this will lead to AI perpetuating biases and impacting vulnerable and under-represented populations. These include the less educated, low skilled, women and elderly, particularly in low and middle-income countries. This is the case because the amount and kind of AI resources are different between developed and developing countries, Organizations, and nations [1, 2]. One way to address this is to provide funding so these areas can begin to close the gap.

On such fund is being proposed by Canada's International Development Research Center. They recommend the formation of a global AI for Development fund, which will establish AI Centers of Excellence in those areas containing vulnerable and underrepresented populations as well as in low and middle-income countries [1, 2]. The goal of the AI Centers of Excellence (CoE) will be to ensure that AI benefits are well distributed and lead to more democratic and open societies. In addition, the AI CoEs AI initiatives will work to ensure that economic gains from AI in societies are widely shared and that no one is left behind [1, 2].

7 Conclusion

Cognitive bias refers to the systematic way in which the context and framing of data, information and knowledge will influence an individuals' judgement and decision making. There are several types of cognitive bias and multiple types of cognitive bias can be influencing your decision making at one time. Biases in how humans make decisions are well documented. Some researchers have highlighted how judges' decisions can be unconsciously influenced by their own personal characteristics, while employers have been shown to grant interviews at different rates to candidates with identical resumes but with names considered to reflect different racial groups. Humans are also prone to misapplying information. For example, employers may review prospective employees' credit histories in ways that can hurt minority groups, even though a definitive link between credit history and on-the-job behavior has not been established.

In many cases, AI can reduce humans' subjective interpretation of data, because machine learning algorithms learn to consider only the variables that improve their predictive accuracy, based on the training data used. In addition, some evidence shows that algorithms can improve decision

making, causing it to become fairer in the process. Therefore, it is important to apply ethical AI practices to remove bias in our AI solutions and the knowledge these solutions provide to drive decision making.

Compliance with ethical standards

Conflict of interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

Reference

1. OECD: OECD recommendation of the council on artificial intelligence, OECD/LEGAL/0449 (2019) Retrieved from <https://legalinstruments.oecd.org>
2. OECD: Artificial intelligence in society. OECD Publishing, Paris (2019). <https://doi.org/10.1787/eedfee77-en>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.