A Framework for Artificial Knowledge Creation in Organizations

Antoine Harfouche, Bernard Quinio, Sana Rouis Skandrani, Rolande Marciniak

► **To cite this version:**

Antoine Harfouche, Bernard Quinio, Sana Rouis Skandrani, Rolande Marciniak. A Framework for Artificial Knowledge Creation in Organizations. ICIS 2017, Dec 2017, Seoul, South Korea. hal-03110617

HAL Id: hal-03110617

https://hal.parisnanterre.fr/hal-03110617v1

Submitted on 14 Jan 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Framework for Artificial Knowledge Creation in Organizations

Antoine HARFOUCHE, Université Francois-Rabelais de Tours
Bernard Quinio, Universite Paris-Nanterre UFR de sciences sociales et administratives
Sana Skandrani, Universite Paris-Nanterre
Rolande Marciniak, IDHES UMR 8533

Introduction

Since the late 1960's, a long overdue attention has focused on artificial intelligence (AI) development for organizational use (Cardenas, 1977). In recent years, artificial intelligence specifically cognitive systems, have surpassed humans in the performance of many tasks. Indeed, among all emerging technologies in the late 20th century, AI provided the most profound impact on decision making in organizations (Lawrence, 1991 ; Shallo & Galliers, 2016). Organizations are in a continuous struggle to make sense of their dynamic business environment. With its ability to avail vast expertise, AI has changed the dynamics of many decision situations (Nonaka, 1994). AI accumulates strategic knowledge that directly guides expert action (Gruber, 1989) and helps to gain new insights to create new knowledge (Shallo & Galliers, 2016). Nonetheless research on AI has primarily focused on technical aspects of AI without explaining how new organizational knowledge may be developed as a result (Arnott & Pervan, 2014).

This paper proposes a new paradigm for managing the recursive aspects of artificial knowledge creation process in organizations. Artificial knowledge is created through a recursive process that extracts, organizes, and aggregates tacit and explicit knowledge in order to articulate, automate, and amplify them. The Framework for Artificial Knowledge Creation explains how knowledge held by individuals in organizations can be simultaneously enlarged and enriched through the recursive amplification of tacit and explicit knowledge enabled by AI.

This paper presents a three-year canonical action research (CAR) project conducted at the CERP (The European Center for Prehistoric Research - Centre Européen de Recherche Prehistorique), a leading French archaeological research center. The CERP conducts research on human history, evolution, and culture and aims to explain changes in human societies through time by reconstructing past lifeways of the Tautavel Man Cave. Since the late 1960s, scientists at CERP have used statistical analysis to abstract the important features from the data collected through surveying and excavation of the Arago Cave. But more emphasis has been placed on gathering information than on effective and efficient use of available knowledge to explain the past behavior of the Tautavel Man Cave. SCHOPPER (Simulation des Comportements des Hommes Préhistoriques dans leur Paléo-Environnement pour la Recherche - Simulation of Prehistoric Men Behaviors in their Paleo-Environment for Research), a three-year canonical action research project, was designed to introduce AI at the CERP center in order to recreate the palaeolithic living conditions at the Arago Cave in Tautavel and consequently, to explain the prehistoric human behavior.

To explore the patterns of AI in organizations, two streams of theory were used. The first follows the seminal work of Nonaka (1994) and Carlile (2002) on Organizational Knowledge: The Dynamic Theory of Organizational Knowledge Creation (Nonaka, 1994) complemented by the Knowledge Boundaries theory (Carlile, 2002) may disambiguate the link between AI use and organizational knowledge creation. The second streams of theory tracks the dominant modern approach of Artificial Intelligence proposed by Russell & Norvig (2010), based on the locus of action in AI while underlining the notion of practical reasoning.

Based on this theoretical underpinning, this paper presents a recursive framework for Artificial Knowledge Creation. This framework offers an analytical perspective on the constituent dimensions of artificial knowledge creation. This framework is composed of five stages: 1) Extracting and Collecting, 2) Curating, 3) Ingesting, 4) Training and testing, and 5) Analyzing and predicting. It is then applied through a clear operational process for facilitating the dynamic creation of appropriate organizational knowledge that can be amplified through the implementation of AI.

Therefore, this paper focuses on how the implementation of an AI agent in organizations is related to organizational knowledge creation. It starts by resuming the different kinds of machine learning used and introducing a taxonomy of Artificial Intelligence Learning Algorithms. Then, it resumes the five stages framework for organizational artificial knowledge creation.

Enlarging organizations knowledge through AI: pre-requisite

Over the past decade, AI has been widely adopted in a number of complex data-intensive fields such as medicine, astronomy, and biology in search of possible solutions to mine the information hidden in the data (Qiu et al. 2016). The use of AI in organizations began with the development of expert systems for problem-solving and decision-making. Artificial intelligence (AI) was defined as: 1) making computers smart, 2) making models of human intelligence, and 3) building machines that simulate human intelligent behavior (Trappl, 1986) in order to aid, and perhaps replace, the decision-makers in organizations (Lawrence, 1991). Whereas, research on AI has been performed mainly in computer science and cognitive psychology (Coulson, 1987). It has primarily focused on expert systems and on the modern machine learning process but without explaining how new organizational knowledge may be obtained as a result (Arnott & Pervan, 2014). This first part of the paper compares AI capacities with the complexity of the knowledge creation in organizations.

Artificial intelligence (Al) in organizations

With machine learning, AI gained the ability to learn without being explicitly programmed. These algorithms can learn patterns in data and make similar patterns predictions in new data while overcoming the need for strictly static algorithms.

Traditionally, there have been three fundamentally different levels of agent supervision in machine learning: supervised, semi-supervised, and unsupervised leaning (Chapelle, et al. 2006). In *supervised learning*, the AI algorithm is given a sequence of desired outputs with the goal of learning to produce the correct output given a new input (Ghahramani, 2004). A concise model is built through a training process in which observations are given with known labels. During this process, the AI algorithm is required to make predictions. Errors are corrected when the algorithm predictions are wrong. The training process continues until reaching a desired level of accuracy on the training data (Kotsiantis, 2007). In *semi-supervised learning*, the AI algorithm uses both labeled and unlabeled data to perform an otherwise supervised learning or unsupervised learning task (Zhu, 2005). The algorithm must learn the structure to organize data as well as to make predictions. In *unsupervised leaning*, the AI algorithm simply receives inputs but obtains neither supervised target outputs, nor rewards from its environment (Ghahramani, 2004). A model is prepared by deducing structures present in the input data to extract general rules. It is done either through a mathematical process that systematically reduces redundancy, or through organizing data by similarity.

In recent years, some of the most impressive advancements in AI have been in the scale of data and the associated learning architectures level of the AI algorithm. Three levels of learning architectures can be considered: Large-scale machine learning, Deep learning, and Deep Reinforcement learning. The outstanding characteristic of these methods is to focus on the idea of learning, rather than on just a single algorithm. The *Large-scale machine learning* scales existing algorithms to work with extremely large data sets (Joachims, 1999). Distributed frameworks with parallel computing are preferred. For example, the Alternating Direction Method of Multipliers (ADMM) has the ability to split or decouple multiple variables in optimization problems, which enables the AI algorithm to find a solution to a large-scale global optimization problem by coordinating solutions to smaller sub-problems (Qiu et al. 2016). Introduced by Hinton and Salakhutdinov (2006), *Deep learning* is based on creating deep neural networks with several hidden layers. The layer closest to the data vectors learns simple features, while the higher layers learn higher-level features (Karhunen, et al. 2015). In contrast to most traditional learning techniques that consider shallow-structured learning architectures, deep learning mainly uses supervised, semi-supervised, or unsupervised strategies in deep architectures to automatically learn hierarchical representations (Qiu et al. 2016). Then, the advent of deep learning has provided reinforcement learning with a “shot in the arm” (Stone et al. 2016). Thus, *Deep Reinforcement learning* enables learning from feedback received through interactions with an external environment (Qiu et al. 2016). The training feedback provided by the external environment constitutes a measure of how well the AI algorithm operates. The algorithm is not told which actions to take, but rather must discover which actions yield the best reward, by trying each action in turn (Kotsiantis, 2007). It shifts the focus to decision making by advancing AI deeper into the realm of learning about and executing actions in the real world.

The Organizational knowledge creation: A complex process

Today, organizations have access to an innumerable amount of information. In this context, they devote a huge effort to find ways to use the available data. The interaction of the Knowledge

Management with Artificial Intelligence makes possible the development of tools that appear as a reply to the expectations to extracts tacit and explicit individual and organizational knowledge in order to articulate and amplify them. This part resumes the seminal work of Nonaka and Carlile on Organizational Knowledge.

Nonaka's Dynamic Theory of organizational knowledge creation

According to Nonaka (1994) « knowledge is created through conversion between tacit and explicit knowledge » and this is through socialization, combination, externalization, and internalization.

Socialization is the process of creating tacit knowledge through shared experience. It is developed through observation, imitation, and practice. On the other hand, Combination is the process of creating explicit knowledge from explicit knowledge. Individuals exchange and combine knowledge through such exchange mechanisms as meetings and telephone conversations. As for the process of converting tacit knowledge into explicit knowledge, Nonaka (1994) coins this as the process of Externalization. Here, metaphors and analogy can be used to enable team members to articulate their own perspectives, and thereby reveal hidden tacit knowledge that is otherwise hard to communicate. On one hand, metaphors help individuals infer from a model of another behavior, by learning through symbols, and by describing the world in terms of prototypes. On the other, analogy reduces ambiguity by allowing the functional operation of new concepts referring to things that are already understood. Reciprocal to externalization, internalization, however is the process of converting explicit knowledge into tacit knowledge; an iterative process of trial and error can trigger internalization through a process of learning by doing.

Furthermore, Nonaka (1994) considers that between deduction, induction, and abduction, it is abduction that has the most import role in the conceptualization process. Indeed, deduction and induction are used when a thought involves the revision of a preexisting concept or the assigning of a new meaning to a concept. Abduction centers on the use of metaphors when there is no adequate expression of an image to create a completely new concept.

Carlile's Knowledge Boundaries Theory

According to Carlile (2002), knowledge is localized and embedded in practice, which makes it difficult to transfer knowledge between actors in multi-disciplinary collaborations. Carlile (2004) proposed a framework that integrates the three perspectives of knowledge boundary, namely, the informational, cultural and political perspectives. The perspective of informational processing assumes that knowledge must be extracted and collected, then effectively transferred through shared lexicons. While, the cultural perspective leverages the practical embeddedness of knowledge, stressing a shared meaning to eliminate the conflicts of interpretation between different actors. Lastly, the political perspective focuses on knowledge transformation by establishing mutual relationships and coordinating interests.

Carlile (2002) proposes three approaches to Knowledge Boundaries that differ in terms of the degree of novelty, dependence and specialization. The three approaches are referred to as syntactic, semantic, and pragmatic. The syntactic boundary may cause difficulties in knowledge transfer as a result of incompatible terminology especially when the degree of novelty, dependence and specialization is low (Carlile, 2004). Therefore, it is important to establish a common lexicon for different actors.

The semantic boundary may cause actors from different disciplines to make different interpretations to the common terminology increasing the degree of novelty, dependence and specialization (Carlile, 2002). Translating interpretations for establishing shared meanings can reduce semantic differences (Carlile 2002, 2004).

The pragmatic approach helps the understanding of consequences that arise between concepts that are different but dependent on each other, even though they might bring novelty to each other. When the degree of novelty, dependence and specialization is high, goal conflicts could emerge. The negative consequences of goal conflicts could be reduced by establishing a common interest in knowledge transformation (Carlile, 2002).

A Framework for Organizational Artificial Knowledge creation

The process of making decisions by reasoning with knowledge is central to AI and to key to successful agent implementation (Russell & Norvig, 2010). This means that representing knowledge is important when implementing an AI algorithm. Knowledge must be extracted and collected into a data base before it can be effectively transferred (Carlile, 2002). According to Storey and Goldstein, (1993), the content and logical structure of the database must be determined. Therefore, the first stage of the organizational knowledge creation must be concerned with the collection of the relevant set of measurements, variables, concepts, and constructs related to the domain and the identification of possible relations between these concepts.

Once knowledge is extracted, a shared meaning must be created to eliminate potential conflicts of interpretation between different actors (Carlile, 2002). Indeed, according to Mykytyn (1990), organizations may try to consolidate the expertise-gathering process by combining many knowledge which could lead to conflicting knowledge acquisition and representation. Therefore, the second stage must establish content curation by sorting, integrating, aggregating, and displaying knowledge in a relevant and usable way.

Data ingestion is performed in the third stage. Even though data preparation could have started during the previous stages, the ingestion of the content is only possible after curation is completed. Data ingestion consists of moving the data into the AI system. It covers a range of steps that deals with the ingestion of structured and unstructured data (Qiu et al. 2016) brought from multiple sources in real time (real-time ingestion), in streaming, or in batch (Assunção, et al., 2014).

After ingestion, the AI algorithm starts learning from a batch of training instances. Preliminary testing and training are considered as part of the fourth stage. During this stage learning will occur by running the algorithm repeatedly through the training set until it finds a prediction vector which is correct on the majority of the training set (Kotsiantis, 2007).

Once general hypotheses are produced, the AI algorithm then can make predictions about future instances. Consequently, this constitutes the fifth and last stage of organizational artificial knowledge creation. During this last stage, prediction rules that were developed during training are used for predicting labels of new data (Kotsiantis, 2007).

AI to Aggregate and Amplify Multi-disciplinary Knowledge in an Organization

The second part of this paper resumes the field of the three-year canonical action research (CAR): The project SCHOPPER.

Canonical Action Research (CAR) for examining artificial knowledge creation in an organization

Canonical Action Research (CAR) is a method frequently employed to conduct empirical research within the IS discipline. Problem identification calls for action and CAR acts as a liberating agent of change (Susman and Evered, 1978; Baskerville, 1999; Davison et al., 2012). While other research methods could have been useful to analyze this process in its natural context, CAR was the most appropriate because of its interventionist approach, which is dedicated to the development of knowledge that is suitable for research and practice (Davison et al. 2012).

The CERP: A multi-disciplinary organization

The CERP is a multi-disciplinary center that works mainly on the Arago Cave at Tautavel. The many multi-disciplinary projects of the CERP rely on surveying, excavation and analysis of data collected since 1964 from processing and recording of archaeological remains of one of the most important prehistoric deposits in the world: The Arago Cave. The project related work draws upon Paleontology, Paleoecology, Paleoclimatology, Zooarchaeology, Paleo-ethnobotany in order to learn about prehistoric societies. Paleontology uses data of the skeletal remains of many individual hominids (such as the Arago XXI, July 1971, and the Arago XLVII, July 1979) to determine human evolution. Paleoecology reconstructs the ecosystem of the past and clarifies the relationship that these homonids and animals had to their environment. Paleoclimatology focuses more on the history of Arago climate and the mechanisms that have changed it. Zooarchaeology studies the faunal remains such as elk, fallow deer, reindeer, musk ox, and bison's bones and shells found at the cave that was occupied from 600,000 to 400,000 B.P. Paleoethnobotany analyses plant remains to identify plants of ancient times, its past environments, climate, ecological, cultural, and human-plant interactions.

Undoubtedly, the process of analyzing archaeological data in such a multi-disciplinary context, involves complex, nonlinear relationships in which cause and effect are not readily distinguished.

Further complicating archaeologists work, the aggregation of a large number of tacit, explicit and mainly unstructured knowledge in a short-time necessitates a large access to resources. As a consequence, addressing these challenges requires both sophisticated modeling and large-scale synthetic algorithms that are only now becoming possible.

The SCHOPPER project 2017-2019

The SCHOPPER project (ANR-DS0701/2016) brings together six partners around a common challenge: The CERP, the CEROS of University Paris Nanterre, Yonsei University Seoul, Catalan Institute of Human Paleoecology and Social Evolution, Craft-ai, and Immersion-Tools. The aim of this collaboration is to develop an innovative AI solution that can recreate the Paleolithic living conditions at the Arago Cave in Tautavel. The "SCHOPPER simulator" will be trained to test (validate or reject) hypotheses related to the prehistoric human behavior in a reconstructed immersive environment. This simulator will be the final result of two interacting digital platforms: the first is based on the archaeologists' database ingested by Craft-ai AI engine. The second is based on archaeologists Data that guides Immersion-Tools to simulate the prehistoric environment.

The SCHOPPER AI will be designed based on the multidisciplinary tacit and explicit knowledge of the CERP researchers. It will use the CERP database that stores 51 years of excavations on a Paleolithic site of world interest: the Arago Cave at Tautavel. A large amount of raw data has been acquired and many archaeometric databases have been developed through the study of the site of Arago as a whole. The methodological approach of using SCHOPPER simulator will be provided by CEROS and will enable to replicate this approach in other contexts. In the past, research hypotheses were often tested by reproducing gestures of prehistoric man, where the limitations of this method were quickly reached. To overcome these limitations, innovative solutions from Machine Learning and virtual reality immersion will be developed as part of this project, to offer new visualization means of scientific results and interaction with the environment.

Research results and discussion

The first outcome of the three-year canonical action research project conducted by the CEROS at the CERP fits the initial objectives expected by the CEROS. Indeed, the first results provided us with an illustration of the dynamic process of transfer of organizational knowledge at the intersection of Nonaka's (1994) theory with the accordance of Knowledge boundaries analysis (Carlile 2002). Figure 1 represents the recursive Framework for Artificial Knowledge Creation with its five stages: 1) extracting and collecting, 2) curating, 3) ingesting, 4) training and testing, 5) analyzing and predicting.

Stage 1: Extracting and collecting the corpus of knowledge

The first stage concerns the extraction and the collection of the "*corpus of knowledge*". Compared to the "body of knowledge" representing the complete set of concepts related to a specific domain, the corpus of knowledge is concerned only with the most relevant. During this first stage, AI experts start to learn the language, the jargon, and the principle thoughts of the domain. Next, they guide the extraction of a rich set of structured, unstructured, tacit, and explicit knowledge. This stage could suggest which fields (attributes or features) of the database are most informative. In the case that this stage is not possible, all the attributes available would have to be measured in order to find and isolate the right informative and relevant features. That is known as the "brute-force" method, which is prone to noise, missing feature values, therefore it is not directly suitable for induction (Zhang et al., 2002). Traditionally, the extraction and the collection have been carried out on an ad hoc basis by a database design expert who had to obtain information about a user's data needs through conducting interviews, examining existing documents and systems, and other such manual, labor-intensive methods (Bouzeghoub, et al., 1985). The main weakness of this traditional approach is that it is usually done by a database design expert who is probably unfamiliar with the specific domain (Storey and Goldstein, 1993). Since the early 90's, AI was able to alleviate these problems with the advocate of expert systems that can generate a database design automatically or semi-automatically. Knowledge-Based Systems for Database Design facilitated the use of heuristics, learning from experience, incorporation of domain-specific knowledge, and the choice of a design strategy. Examples of AI Knowledge-Based systems used for this purpose are: Intelligent Interview Systems (I2S), Computer Aided Requirements Synthesis (CARS), and Modeller (Storey and Goldstein, 1993).

In the case of SCHOPPER project, because of the multidisciplinary nature of tacit and explicit knowledge used by actors, a combination of an ad hoc intervention of database design expert with the use of a knowledge based system is necessary, aided by the exploitation of mind mapping techniques.

Researchers from different disciplines at CERP were invited to *externalize* their knowledge using Vu software, assisted by two database design experts. This method was used to extract and structure CERP researchers' tacit and explicit knowledge. The two experts obtained information about knowledge used by decision makers through interviewing and by examining existing published reports.

Stage 2: Curating the content

The second phase concerns the process of gathering, aggregating, organizing, and sorting individual knowledge within a vast volume of consolidated knowledge and presenting it in an organized and meaningful format. During this stage, the organizational knowledge is sorted, aggregated, and displayed in a relevant and usable way. During data preparation and curation, researchers choose methods to handle missing data (Batista & Monard, 2003), to handle outlier (or noise) detection (Hodge & Austin, 2004). They must identify and remove as many irrelevant and redundant features as possible (Yu & Liu, 2004).

In the case of SCHOPPER, many activities occurred during the process of content curation: First, database design experts started by the aggregation of knowledge that is relevant for explanation and prediction. Second, they had to deal with the distillation of knowledge in a simplified format that shares only the most relevant and important knowledge. Third, they executed many mashups, which creates outcome resulting from the merger of many curated content. Fourth, they finished with the elevation, which identifies the significant trends that has emerged from exchange of knowledge. The combination of individual knowledge helped creating a collective intelligence while reducing dimensionalities of the data. This enables the AI algorithm to operate faster and more effectively. Indeed, the fact that many features in the database were correlated, this unduly influenced the accuracy of the process. This problem has been addressed by constructing new features from the basic feature set. This technique is known theoretically as « the feature transformation » (Markovitch & Rosenstein, 2002).

Stage 3: Ingesting

Content ingestion allows the automated ingestion of assets and metadata from the database to the Artificial Intelligence platform. Data ingestion focuses on moving large amounts of curated data into the AI platform (Assunção, et al., 2014). The ingestion may be done in real time (real-time ingestion), in streaming, or in batch (Assunção, et al., 2014). Its architecture must be flexible based on streaming data flows with a failover and recovery mechanism. It either actively polls for data or passively waits for data to be delivered to the AI platform.

In the case of SCHOPPER, data and metadata were passed to the AI platform at runtime bases in-batch learning fashion with the decision to collect the full knowledge before starting the training process. The in-batch ingestion increased the control of the transformation logic realized through multiple data ingestion, allowing the operationalization of many predictive models.

Stage 4: Training and testing

Training is the abductive part of the process that will let the AI algorithm to learn a set of rules from instances creating a classifier that can be used to generalize from new instances. The AI system learns by running the algorithm repeatedly through the training set until it finds a prediction vector which is correct on the majority of the training set (Kotsiantis, 2007). The prediction accuracy of the AI algorithm is measured by dividing the number of correct prediction by the total number of predictions.

In the case of SCHOPPER, as the problem was complex, the decision was to construct several subsystems that use different machine learning strategies to gain power to analyze the questions and arrive at the most likely answer. The technique that was used to train the AI algorithm was to split the training set by using a part for training and the other part for estimating the prediction accuracy. When the error rate was unsatisfactory, database design experts returned to previous stages 1, 2, and 3 because it meant that relevant features for the problem were not fully captured. This recursive process helped identifying the most relevant feature in an abductive way confirming the importance giving by Nonaka to abduction. To follow the training of the AI, the CERP key actors created a self-organizing team and organized weekly 'Reflexivity Sessions' to conduct a meaningful dialogue, to think about the tool, and to interact with the simulation result. The 'Reflexivity Sessions' established a recursive and reflexive relation between individual knowledge and the organizational artificial knowledge.

Stage 5: Analysis & Prediction

Prediction rules developed during the training stage are used to predict the labels of new data. The feature that best divides the training data became the root node of a decision tree. Indeed, a decision tree is created to classify instances by sorting them based on feature values (Kotsiantis, 2007). Each node represents a feature in an instance to be classified, and each branch represents a value that the node can assume. The decision tree makes the analysis comprehensible. Decision makers can easily understand which instance belongs to which class.

In the case of SCHOPPER, the crucial point was to determine how to decode the predictions of the classifiers for a final prediction.

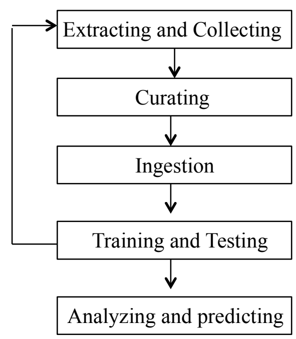

Figure 2. The Five stages of the Framework for Artificial Knowledge Creation in Organizations

Conclusion

This paper aims to demonstrate that AI applications in organizations are simply an array of new extension tools for knowledge management. Instead of simply presenting how managers are using the AI for supporting decisions, we have focused on how AI extracts, organizes, and aggregates tacit and explicit knowledge in order to articulate, automate, and amplify them.

The problem-solving approach that was used focuses on insights that can be induced from problem-solving activities. It starts with the application of an intervention plan in order to solve the diagnosed problems in the target organization. This intervention plan is guided by a theoretical perspective: Nonaka's Dynamic theory of organizational knowledge creation and Carlile's Knowledge Boundaries theory. While the former represents the focal theory, it indicates the kind of transfer of knowledge that will occur. The latter, indicates how the knowledge will transfer and will be created. After the application of the intervention plan, researchers used data generated from the problem-solving activities to compare and contrast with existing IS theories, and to develop new theoretical knowledge. The results can be resumed in a Framework for Artificial Knowledge Creation with its five stages: 1) extracting and collecting, 2) curating, 3) ingesting, 4) training and testing, 5) analyzing and predicting.

The canonical action research conducted, has identified two main practices that were triggered by the development of the AI agent: 1) the ability to initiate a dialogue between the different actors which can lead to a consolidation of the organizational knowledge, and 2) the ability to establish recursive and reflexive relation between individual knowledge and the organizational artificial knowledge.

This paper presents a framework that underpins how AI algorithms can develop organizational artificial knowledge using abductive-ly predictive models. The limitation is that the framework cannot demonstrate causality, but it could enable researchers to better estimate the causal effect of a treatment.

References

Arnott, D., and Pervan, G. 2014. “A critical analysis of decision support systems research revisited: the rise of design science.” *Journal of Information Technology*, (29), pp. 269–293.

Assunção, M. D., Calheiros, R. N., Bianchi, S., Netto, M.A.S., and Buyya, R. 2014. “Big data computing and clouds: trends and future directions,” *Journal of Parallel and Distributed Computing*, (79-80), pp. 3–15.

Baskerville, R.L., 1999. “Investigating information systems with action research.” *Communications of AIS*, (2:3), pp. 1-32.

Batista, G., and Monard, M.C. 2003. An Analysis of Four Missing Data Treatment Methods for Supervised Learning, *Applied Artificial Intelligence*, (17), pp. 519-533.

Cardenas, A.F. 1977. Technology for Automatic Generation of Application Programs -A Pragmatic View, *MIS Quarterly*, (1:3), pp. 49-72.

Carlile, P. R. 2002. “A Pragmatic View of Knowledge and Boundaries: Boundary Objects in New Product Development,” *Organization Science*, (13:4), pp. 442-455.

Carlile, PR. 2004. “Transferring, translating, and transforming: An integrative framework for managing knowledge across boundaries.” *Organ. Sci.* (15:5), pp. 555–568.

Chapelle, O., Zien, A., and Scholkopf, B. (Eds.). 2006. *Semi-supervised learning*, MIT Press.

Coulson, R. N., Folse, J. L., and Loh, D.K. 1987. “Artificial intelligence and natural resource management,” *Science*, (237), pp. 262-267.

Davison, R. M.; Martinsons, M. G,; and Ou, C. X.J. 2012. "The Roles of Theory in Canonical Action Research," *MIS Quarterly*, (36: 3) pp.763-786.

Evermann, J. 2005. Towards a cognitive foundation for knowledge representation, *Info Systems J*, (15), 147–178.

Ghahramani, Z. 2004. “Unsupervised Learning,” in *Advanced Lectures on Machine Learning*, Bousquet et al., (Eds.), Springer Verlag.

Gruber, T. R. 1989. *The Acquisition of Strategic Knowledge*, San Diego, CA: Academic Press.

Hinton, G. and Salakhutdinov, R. 2006. Reducing the dimensionality of data with neural networks. *Science*, (313), 504-507

Hodge, V., and Austin, J., 2004. A Survey of Outlier Detection Methodologies, *Artificial Intelligence Review*, (22:2), pp. 85-126.

Joachims. T. 1999. “Making large-scale support vector machine learning practical.” In *Advances in kernel methods: support vector learning*, B. Scholkopf, C. J. C. Burges, and A. J. Smola, (eds), pp. 169-184. MIT Press, Cambridge, MA, USA.

Karhunen, J, Tapani R, and KyungHyun C., 2015. "Unsupervised Deep Learning: A Short Review." *Advances in Independent Component Analysis and Learning Machines*, (125). pp.

Kotsiantis, S. B. 2007. “Supervised machine learning: A review of classification techniques.” *informatica* (31), pp. 249-268.

Lawrence, T., 1991. “Impacts of artificial intelligence on organizational decision-making.” *Journal of Behavioral Decision Making*, (4 :3), pp. 195-214.

Markovitch, S. & Rosenstein D. (2002), Feature Generation Using General Construction Functions, *Machine Learning*, (49), pp. 59-98.

Mykytyn, K., Mykytyn, P. P. Jr., and Slinkman, C.W. 1990. Expert Systems: A Question of Liability?, *MIS Quarterly*, (14: 1), pp. 27-42.

Nonaka, I., 1994, “A Dynamic Theory of Organizational Knowledge Creation,” *Orgnization Science*, (5:1). pp. 14-37.

Qiu, J., Q. Wu, G. Ding, Y. Xu, and S. Feng, 2016. “A survey of machine learning for big data processing,” *EURASIP Journal on Advances in Signal Processing*, (1), pp. 1–16,

Russell S. & Norvig P. 2010. *Artificial Intelligence, A Modern Approach*. Third Edition. Pearson Education.

Shah, J., Tambe, M. and Teller, A. 2016. " Artificial Intelligence and Life in 2030," One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA, September 2016, http://ai100.stanford.edu/2016-report.

Shallo, A. and Galliers, R. D., 2016. “Towards an understanding of the role of business intelligence systems in organisational knowing,” *Information Systems Journal*, (26:4), pp. 339-367.

Stone, P., Brooks, R., Brynjolfsson, E., Calo, R., Etzioni, O., Hager, G., Hirschberg, J., Kalyanakrishnan, S., Kamar, E., Kraus, S., Leyton-Brown, K., Parkes, D., Press, W., Saxenian, A., Shah, J., Tambe, M., Teller, A. (2016): Articial Intelligence and Life in 2030. Tech. rep., One Hundred Year Study on Articial Intelligence: Report of the 2015-2016 Study Panel, Stanford University, Stanford, CA, http://ai100.stanford.edu/2016-report

Storey, V. C. and Goldstein, R. C., 1993, Knowledge-Based Approaches to Database Design, *MIS Quarterly*, (17: 1), pp. 25-46

Susman, G.I., Evered, R.D., 1978. "An Assessment of the Scientific Merits of Action Research", *Administrative Science Quarterly,* (23), pp. 582-603.

Trappl, R. 1986, "Impacts of artificial intelligence: An overview', in Impacts of Artificial Intelligence, R. Trappt (ed.). Amsterdam: North-Holland.

Tsoukas, H. 2009. A dialogical approach to the creation of the new knowledge in organizations. *Organizational Science*, (20:6), pp. 941-957.

Yu, L., Liu, H. 2004. Efficient Feature Selection via Analysis of Relevance and Redundancy, *JMLR*, (5:4), pp. 1205-1224.

Zhang, S., Zhang, C., Yang, Q. 2002. Data Preparation for Data Mining. Applied Artificial Intelligence, Volume 17, pp. 375 - 381.

Zhu., X. 2005. Semi-supervised learning literature survey. Computer Sciences Technical Report 1530, University