The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry

Daniel Viberg and Mohammad H. Eslami

> *"The ambiguity, as a consequence of knowledge integration becoming a fashion word, is in itself an explanation for why theorists sometimes end up in incomplete explanation."*
>
> Anna Jonsson
> (Translated from Swedish)

The impact of such current state-of-the-art technology as machine learning (ML) on organizational knowledge integration is indisputable. This paper synergizes investigations of knowledge integration and ML in technologically advanced and innovative companies, in order to elucidate the value of these approaches to organizational performance. The analyses are based on the premise that, to fully benefit from the latest technological advances, entity interpretation is essential to fully define what has been learned. Findings yielded by a single case study involving one technological firm indicate that tacit and explicit knowledge integration can occur simultaneously using ML, when a data analysis method is applied to transcribe spoken words. Although the main contribution of this study stems from the greater understanding of the applicability of machine learning in organizational contexts, general recommendations for use of this analytical method to facilitate integration of tacit and explicit knowledge are also provided.

Introduction

The rapid pace of innovation in the context of new technology development has attracted significant attention of technology firms, as this offers potential for using these tools for knowledge integration as a means of creating and sustaining competitive advantage (Grant & Baden-Fuller, 2004). Previous research has shown that knowledge integration has great potential to accelerate innovation, since identifying and combining distributed knowledge can enhance the competitive advantage of firms, distinguishing them from their competitors (Carlile & Rebentisch, 2003; Yang, 2005). However, while firms acknowledge the advantages and necessity of knowledge integration, they typically face different difficulties in accessing distributed knowledge (Enberg et al., 2006; Schmickl & Kieser, 2008). In other words, integrating distributed knowledge is challenging, especially tacit knowledge, as this is knowledge gained through personal experience, making it difficult to transfer or codify.

Despite these difficulties, many researchers are of the view that artificial intelligence (AI), and machine learning (ML) in particular, can be adopted in organizational knowledge integration (Li & Herd, 2017). Paradoxically, ML algorithms rely on experience-based knowledge (Jin et al., 2018), in the sense that large sets of data are interpreted in order for some general rules to be codified. In this context, it is important to study firms that focus on technology-based activities (Berggren et al., 2011), because their innovation efforts are under the effect of two major influences: (a) strong interrelationship activities between R&D and production, and (b) changes in the character of AI that can mandate either specialization or increase in complexity (Lin & Chen, 2006). These two trends necessitate further exploration of AI for knowledge integration activities within organizations.

Thus far, this topic has not been sufficiently investigated, as AI, and ML in particular, are relatively recent innovations. Thus, the aim of the present study is to investigate how ML could be combined with human ability and knowledge in a broad sense in order to help in the acquisition and transfer of different types of

The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry *Daniel Viberg and Mohammad H. Eslami*

knowledge. To achieve this goal, the following research question, pertaining to the fundamentals of both domains, is addressed in the present study: *How can ML facilitate tacit and explicit knowledge integration in technological organizations?*

By answering this question, the study contributes to both individual and organizational knowledge, while expanding the current scholarship on links between AI and knowledge integration theories. The next section provides a review of extant literature, focusing on knowledge integration and ML. This is followed by a brief description of the research design adopted in this study, after which the results are presented. The paper concludes with a discussion of key findings and their implications for research and practice.

Literature Review

Knowledge integration
According to Grant (1996a), knowledge integration is "a process for coordinating the specialized knowledge of individuals", whereas other researchers define knowledge integration as a combination of various activities. For instance, Tell (2011) distinguished among studies in which knowledge integration is defined as transferring or sharing knowledge (Huang & Newell, 2003; Marsh & Stock, 2006), applying similar/related knowledge (Teece et al., 1997), and combining specialized and complementary knowledge (Kogut & Zander, 1992; Tiwana & McLean, 2005). However, according to Okhuyen and Eisenhardt (2002), the concept of knowledge integration exceeds sharing and transferring to include combining specialized knowledge in order to create new knowledge.

While there is presently no consensus on the definition of knowledge integration, most researchers make a clear distinction between explicit and tacit knowledge (Nonaka & Takeuchi, 1995; Grant, 1996b; Spender, 1996). Explicit knowledge refers to factual knowledge that can be easily recorded in, for example, manuals, written policies, and procedures (Ernst & Kim, 2002). However, engineering knowledge– such as that based on experience, intuition, and professional judgment– is tacit (Backlund, 2006). As this is the greatest source of innovation, companies must continually regenerate and capture their tacit knowledge (Grant, 1996b; Hansen et al., 1999). If a company fails to retain the tacit knowledge of its most experienced and talented employees, then its overall competence will falter. Therefore, allowing employees to learn from each other

is a key requirement for companies' long-term success (Argote & Miron-Spektor, 2011). Tacit knowledge is not a static stock of knowledge, as it is continuously expanding through deliberate and experiential learning. However, it is also rapidly eroding, as it relies on the memory of those that possess it. The degree of tacitness can also vary; the less explicit and codified knowledge tends to be, the more difficult it will be for individuals and organizations to assimilate it (Howells, 1996). The knowledge a person possesses can also require application of explicit theories to practical situations or problems.

In most cases, knowledge with both explicit and tacit elements is required, especially when performing a development task. Hence, as Jonsson (2012) argued, it is important to know to what degree some knowledge is explicit and/or tacit, as this will help define appropriate methods for transferring and translating this knowledge. To aid in this process, nearly 30 years ago, Nonaka (1994) developed a model depicting four stages of individual knowledge creation, denoted as Socialization, Externalization, Internalization, and Combination. The initial process of Socialization involves exchanging exclusively tacit knowledge between individuals through activities like master-to-apprentice mentorship or informal meetups. Externalization relates to articulating the tacit knowledge to explicit notions, and can be performed using various mediums, such as metaphors and stories. Combination involves combining explicit knowledge in different ways, which can be done via IT systems or knowledge banks. Finally, during Internalization, explicit knowledge is ingrained into corporate culture and work methods. In sum, this entire sequence modifies the tacit knowledge that an individual currently holds, by introducing new explicit knowledge (Jonsson, 2012; Nonaka, 1994; Wiig, 1997).

Machine learning (ML)
ML is an analytical technique whereby an algorithm is developed based on computational statistics derived from available data. The main goal of ML is to apply rules developed through exposure of large datasets to new (but similar) scenarios (Witten et al., 2016). The process, in its simplest form, relies on classification, whereby a dog can be classified as an animal or a dog, depending on the level of specificity, or a voice input can be recognized as language and thus translated into text. The ML field experienced an initial expansion in 1983, when researchers started asking questions related to how, what and why machines should learn (Simon, 1983; Provost, 2000; Feldman, 2011).

The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry *Daniel Viberg and Mohammad H. Eslami*

**Table 1.** An estimated square function

Input	1	2	3	...	n
Output	1	4	9	...	n^2

It has since advanced considerably, giving rise to various types of ML, one of which is supervised learning, which involves taking labelled datasets, learning from them, and then labelling new datasets. An example of this is function estimation based on a set of input and output numbers. If a learning algorithm is told that the first number in a set, which is 1 in the case described in Table 1, produces 1 as the output, 2 yields 4, 3 relates to 9, etc., it can deduce that the function relating input to output is x^2. It is important to note that the algorithm will always provide only an estimation, which will then become more accurate as more data becomes available (Kubat, 2015).

Supervised learning is based on induction, as the algorithm takes a set of examples and tries to extrapolate those examples to some generalized rule. For example, if the question "Did the sun rise during the last 10 days?" produces a "yes" response, then it would deduce that the sun will rise tomorrow as well. Figure 1 shows a simplification of supervised learning.

In contrast to supervised learning, unsupervised learning is comprised of large sets of data that are provided to the system from which structure is created based on the relationships among input parameters (Kubat, 2015). In other words, the algorithm is responsible for assigning

**Figure 1.** Supervised learning labels data well
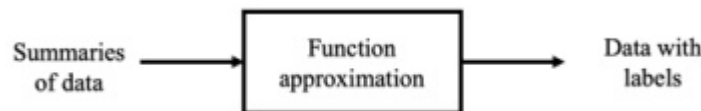
meaning to input data. For example, an algorithm can be given a large set of images of different dogs and will learn to conclude that a picture of a car is not the same as those previously shown. Another example relates to crowd classification (Hoffer & Ailon, 2015). Given a large crowd, the algorithm will learn to distinguish males from females, individuals with facial hair from those without, people of different ethnicities, etc. (see Figure 2 for a simplification of the algorithm). Still, as is the case with supervised learning, larger data sets serve to make unsupervised learning algorithms more accurate (Kubat, 2015).

Clustering and applicability of ML
Clustering is one method of summarizing collected

data. It can be of hard and soft type and entails collating data points into groups based on some measure of similarly (Kearns et al., 1998). In the hard-clustering approach, classification is binary, whereby the input data points either belong or do not belong to a group, whereas soft clustering assigns probability of belonging to a group. In practice, hierarchical clustering is often applied, as it allows building hierarchal clusters of data groups with applications in recommendation engines, market segmentation, social network analysis, search results grouping, medical imaging, image segmentation, or anomaly detection (Domingos, 2012). A particularly useful feature of hierarchical clustering algorithms is their ability to handle multidimensional data, which are usually utilized in ML (Dugad & Ahuja, 1998).

Figure 2. Unsupervised learning clusters data well
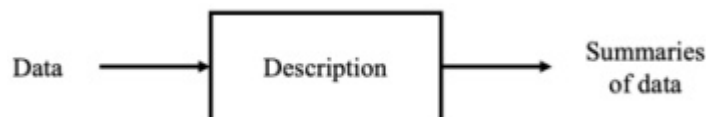
The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry *Daniel Viberg and Mohammad H. Eslami*

According to Hodson (2016), in order to ascertain if ML can be applied to a certain context, the problem at hand should be examined, followed by the available data, as well as feasibility and expectations of the ML process. It is also essential to understand the difference between automation and learning problems, as ML can facilitate automation, whereas not all problems require learning ability. In practice, automation without learning can be applied to scenarios when predefined sequences of steps– typically executed by humans– are consistent, and are executed in a similar manner, and therefore do not require any flexibility in a problem-solving algorithm (Hodson, 2016)

Problems that require automation paired with learning typically, (1) involve prediction rather than causal decision making (in other words, the average relatability in data is of interest), or (2) are sufficiently self-contained, or relatively insulated from outside influences, as this would allow the algorithm to make relevant inferences (even though it will not be able to learn anything beyond the data provided) (Hodson, 2016). It is also important that the data provided has certain characteristics. Hall et al. (2016) outlined a few aspects that ML data should fulfil to be valid, namely, that the data should not be biased, or contain any misleading information or missing values, as more reliable data would yield a more precise algorithm. In summary, ML is applicable when: (1) there is a problem that requires prediction rather than causal interference, (2) the problem is insulated from outside influences, (3) a large dataset is available for training, (4) the training data does not contain misleading information, and (5) the training data is not biased (Hodson, 2016).

Research Design

To address the research question guiding this investigation, a single case study at a large technological firm in Sweden was conducted. The company employs more than 15,000 employees and has R&D centres at different locations where an extensive variety of high technologies with applications in different industrial areas are developed, along with industrial products and services. The company applies several approaches toward product development projects, whereby R&D engineers collaborate with other units to identify key features of new products and services, as well as potential challenges that may arise in their design and production.

Given that the focus of the present investigation was on use of ML, three specific functions within the organization, namely design, industrialization, and production, were chosen as the research object. The design division is responsible for developing detailed models of parts comprising the final product, whereas the industrialization division ensures that those parts can be manufactured in practice, and the production division is tasked with physically creating these parts. These three divisions are in constant collaboration with each other via digital and physical interfaces that allow them to share pertinent knowledge. Our observations at the site revealed that some of the collaboration took the form of standard documents that were used by all parties. The information pipeline is of particular interest for the present study because the company leadership suspected that the manner in which both explicit and tacit knowledge was used and shared was inefficient.

Based on information that emerged during the discussions with representatives of all the aforementioned units, it was surmised that both knowledge integration and ML could address the efficiency issues. As discussions progressed, a specific research question emerged: *Can both explicit and tacit knowledge be handled by ML, with the aim of more efficient knowledge integration?*

As our aim was to provide a knowledge integration solution for a specific technology company, as well as to relate that outcome to ML theories, it was important to obtain as many opinions on the subject as possible, as this would allow us to elucidate how consistent interpretations of knowledge integration were throughout the company. Moreover, by analyzing this information, a more precise practical definition of knowledge integration could be adopted in the study.

To meet the study objectives, three sources of data were used: (1) relevant documents that contained an explicit form of knowledge integration as well as a strategy for tacit knowledge integration; (2) formal and informal interviews, focusing primarily on the tacit aspects of knowledge integration; and (3) the so called "Go to Gemba" strategy, based on learning theory and its related philosophy (Liker & Meier, 2006). The last data collection method required assessing the way systems operated, documents were generated, and meetings were conducted within the three hierarchal levels and divisions. The investigation was deliberately not limited to documents and systems only, since decisions pertaining to how and when meetings should be performed could influence tacit knowledge integration.

The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry *Daniel Viberg and Mohammad H. Eslami*

Table 2. The interview strategy

Category	Limitation
Structure	Semi-structured interviews guided by predefined questions that allow for further discussions
Size	3 (hierarchical) ×3 (divisional) = 9 actors
Communication medium	Face-to-face
Responsibility	Management Group leader First line subordinates

Moreover, there were also some indications that official procedures and guidelines were not always followed. In other words, the explicit description of how tacit knowledge should be shared and documented likely deviated from actual methods adopted in practice. To facilitate our case analysis, a significant amount of information about the company's structure, culture, strategy, and current knowledge integration methods was obtained via informal discussions with relevant company staff who were knowledgeable on the topic of knowledge integration.

In addition, as noted earlier, formal interviews were conducted with key informants. The primary objective of individual interviews was to, (1) elicit the employee's interpretation of knowledge integration and usage for retrieving relevant information; (2) establish what the employees would like to do and have in order to develop their knowledge; and (3) identify potential areas where ML could be of use. To aid in addressing the third aim, pertinent literature on ML was examined.

Analysis

Based on findings yielded by an analysis of all pertinent information, ML might be used for increasing workplace efficiency, as the resulting automation of administrative and repetitive tasks can save managers time. Moreover, ML can benefit company staff by assisting them with finding the most relevant knowledge sources on topics of interest. Based on these assertions, a general description of how ML fits in the knowledge integration perspective is suggested using knowledge integration theories, with an aim of specifying data suitable for ML applications, as well as how it can be sourced and used.

However, it is important to note that ML is not suited for determining how to reach value-dependent goals (Domingos, 2012; Kubat, 2015). This is intuitive, as an algorithm does not "understand" what it is learning, as the value of the knowledge it entails is not considered during the learning process. Consequently, when data is biased or erroneous, a ML algorithm would produce incorrect output. Hence, to fully benefit from ML in practice, a value interpreting entity (usually a human expert) is required to provide feedback to the ML algorithm during the learning phase.

In sum, ML is a tool that will not function without the assistance of a value interpreting entity. Such algorithms can utilize both explicit and (some types of) of tacit knowledge. Explicit forms of knowledge are codifiable, objective, not connected to a specific context, simple to transfer, and are often described as "data" or "information". Typical forms of explicit knowledge that ML algorithms use involve numerical data, images, and transcriptions, that is, objective input, since correlations among subjective data are very difficult for an algorithm to "interpret". For example, if an image recognition algorithm was used to find pictures of objects and the

The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry *Daniel Viberg and Mohammad H. Eslami*
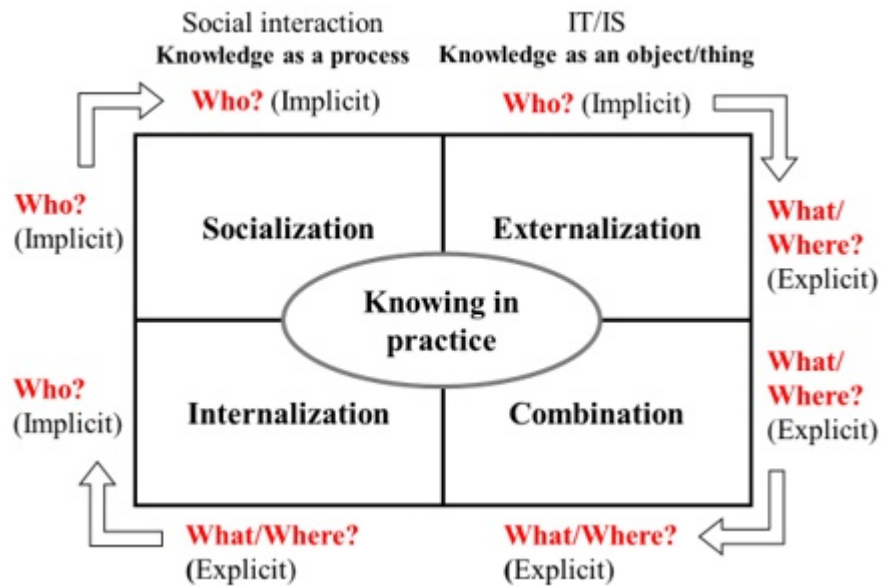
Figure 3. Knowledge sourcing issue represented in the SECI model, adapted from Jonsson (2012)

user searched for a car, it would be successful in providing an image of a car. However, if the user decided to search for an attractive-looking car, the algorithm would not be able to provide an image that would meet this requirement, since attractiveness is subjective. Nonetheless, it could correlate large amounts of data to determine what most users considered an attractive-looking car.

This last example indicates that ML algorithms can acquire a certain level of tacit knowledge, even though their inferences may not fully correspond to a particular user's interpreted preferences. Thus, based on the absence of value interpretability in ML, there is a limit to what types of tacit knowledge an algorithm can acquire. For example, in a scenario where a ML algorithm learns to identify voices and attribute them to individual speakers, the algorithm will work as intended through experience, gaining an ability that is related to the tacit form of knowledge. Still, that experience is relevant only to a specific case, because the algorithm would not be able to label unfamiliar voices or perform any other function for which it was not specifically trained. In

summary, ML is expected to handle explicit knowledge well, along with a certain level of tacit knowledge.

Discussion

As mentioned earlier, the purpose of this study was to investigate how ML technology can facilitate tacit and explicit knowledge integration in technological organizations. The key issue that emerged during the investigation pertained to the difficulty of identifying either explicit information or a person with the right knowledge. Since ML involves learning by experience, it is interesting to investigate if issues related to recognising individuals' knowledge domains– one of the mechanisms lacking in "common knowledge"– and knowledge sourcing in documents, could be improved. In this section, therefore, a speculative framework of how ML could be used to solve this issue is presented based on the SECI model (Nonaka, 1994). As shown in Figure 3, during the different phases of knowledge creation, several different mediums are used. For example, during the socialization phase, employees talk to each other to obtain the required information, whereas documents
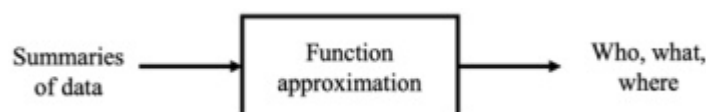
Figure 4. A supervised learning algorithm in the smart knowledge bank

The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry *Daniel Viberg and Mohammad H. Eslami*
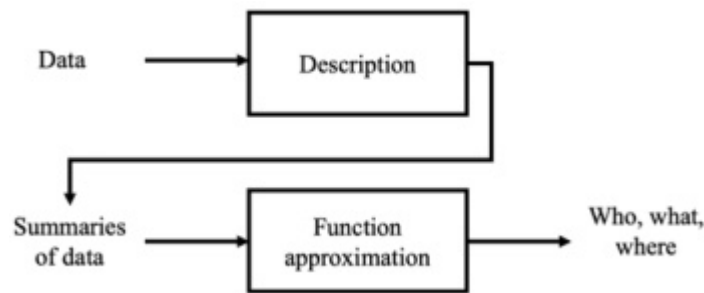
Figure 5. A smart knowledge bank algorithm using unsupervised and supervised learning

and manuals are generated in the combination phase. Thus, the aim was to ascertain who possessed what types of knowledge and where documents or manuals were stored.

As stated previously, in the company where our case study was conducted, there seemed to be a correlation between knowing where to find relevant information and the employee's duration of employment at the company, which was also related to the size of an individual's social network within the firm (Nonaka, 1994; Miller & Fern, 2007). From the perspective of ML, the question thus became how to retrieve and share the ability of experienced staff to identify the most likely individual that possesses knowledge on any given question or topic, or determine where that knowledge

was stored, as well as if it was shared with any other individual. A solution could then be constructed from the initially determined output, which in this case would be a classification that results in labelled data that contains the knowledge source, such as a person, document, or system. Moreover, input required for the ML algorithm to reach a decision would be derived from the question being asked, which could be related to a broad topic, such as "thermodynamics", or a very specific subject, such as "entropy." The resulting algorithm could function as shown in Figure 4.

Summaries of data would in this case consist of data clusters that contain roles, assignments, documents, manuals, systems, topics, projects, people, divisions, etc. Clearly, clustered data with such dimensions, however,

Table 3. Data dimensions required for a smart knowledge bank algorithm

| Data dimensions	Collection method
Role	Documented digitally
Division	Documented digitally
Assignment	Documented digitally
Project	Documented digitally
Systems used	Traced digitally
Documents used	Traced digitally
Manuals used	Traced digitally
Colleagues spoken to	Traced using ML transcription
Alleged knowledge	Traced during meetings + mail and chat conversations

The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry *Daniel Viberg and Mohammad H. Eslami*

would be difficult to collect and modify manually (Karypis et al., 1999). Therefore, an unsupervised learning algorithm should be employed to identify any clusters in the various data dimensions. In this case, it would involve correlating data points, that is, a particular system with a topic, or a specific person with a role. The type of algorithm that would be suitable in this case is based on a hierarchical clustering method because it does not necessitate prior knowledge of the number of clusters required, which is not the case for other methods. Additionally, hierarchical clustering is suitable because this scenario entails relatively low data quantity (Kubat, 2015). An augmented algorithm is shown in Figure 5.

Having the right data is a prerequisite for developing an algorithm as exemplified in Figure 5. Some of the data requirements are presented in Table 3, where the data dimension represents types of information collected, while the collection method describes how the information is to be collected. For example, the Role, Assignment, or Project related to a particular employee should be obtained from digital documents. On the other hand, the Systems used, or Documents used by an employee, would be traced, since it is not initially possible to know which documents someone has read. Colleagues spoken to– which also includes the discussion subject– would be gathered using an appropriate ML method.

Using this strategy, data pertaining to all employees, projects, documents, and manuals, their usage and knowledge shared, would be hierarchically clustered. Speculatively, it would be possible to gradually arrive at an algorithmic output that pertains to what, where, or who holds knowledge that is connected to a specific topic, assignment, or question. On the other hand, ML cannot be used to classify knowledge, since such an algorithm would be highly dependent on the result of the clustering and classification's interpretive reliability. A representation of the expected process' output is shown in Figure 6.

This proposition for ML data clustering was tested using the Google visualization tool known as tensor flow, as shown in Figure 6. Tensor flow takes vectors in high-dimensional space and visualizes them (plots them visually) along with their correlations with other vectors in a lower-dimensional space. Using this tool and a dataset that simulates employees' knowledge and experiences related to various tasks, documents, projects, systems, and hierarchical levels, it was possible to generate a visual cluster representing the output of a hierarchical clustering algorithm. The data utilized in this scenario consisted of 3,000 rows (each row representing a person) and 27 columns, each of which represented the time an employee spent on a certain task, system, project, or document. The graph in Figure 6 shows how all employees (denoted as dots) correlated with other, similar employees.
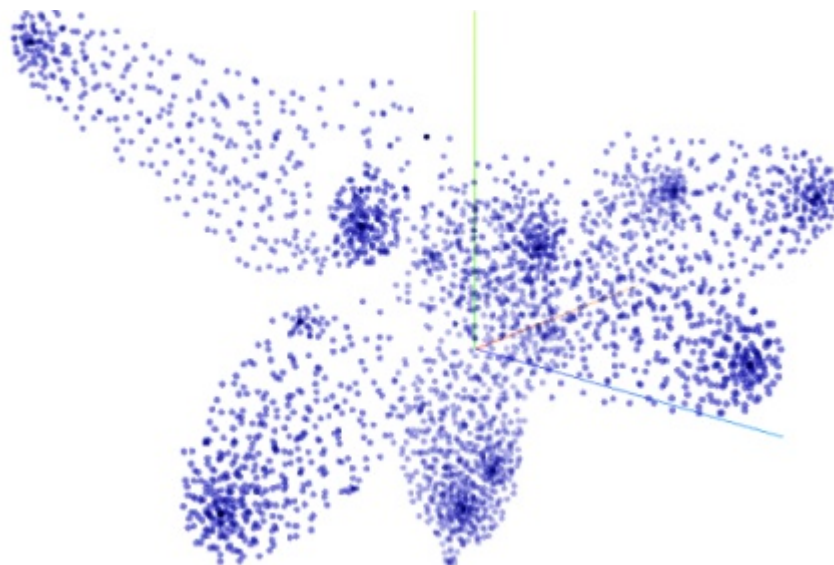
Figure 6. A visualization of how a smart knowledge bank could cluster experiences into various categories

The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry *Daniel Viberg and Mohammad H. Eslami*

In this context, "similar" means that the employees are at the same hierarchal level, embody the same experiences in a system, or have spent the same amount of time looking at a document. The branches represent groups of individuals who are connected to a specific task, document, or other relevant differentiating aspect, such as a project or role. Although ML seems promising theoretically, there are many obstacles that must be overcome in order to implement such a system in practice. Some of these are related to various limitations, namely hardware, software, or performance, as well as ethical issues and interpreting organizational culture.

Conclusion

Theoretical contribution
The aim of this investigation was to provide insight into the role of ML in tacit and explicit knowledge integration in technology firms. The findings presented here contribute to the ongoing debate on the value of knowledge sharing and integration within organizations in generating competitive advantage. The case study results suggest that ML technology cannot be viewed as a static knowledge bank, like typical IT systems. Moreover, as some forms of tacit knowledge can be interpreted by ML algorithms, it would be beneficial to revise the SECI model proposed by Jonsson (2012).

ML can also be applied to capture verbal communications and knowledge exchange during meetings, as the content can be transcribed based on supervised learning, and then stored for others to read. In addition, ML provides companies with the ability to automate the way that individual employees search for knowledge and store information that others might find useful.

Practical implications
As knowledge is becoming increasingly significant in organizational activities, managers need to consider adopting AI for integrating tacit and explicit knowledge. This is particularly important for firms that rely on innovation and R&D. The findings reported in this work could aid companies in determining what types of knowledge they wish to integrate and how best to utilize ML in this process

References

Argote, L., and Miron-Spektor, E. 2011. Organizational learning: From experience to knowledge. *Organization science*, 22(5): 1123-1137.

Backlund, G., and Sjunnesson, J. 2006. Better systems engineering with dialogue. *Dialogue, Skill and Tacit Knowledge*: 135-151.

Berggren C., Bergek A., Bengtsson L., So derlund J. 2011. *Exploring knowledge integration and innovation. In Knowledge integration and innovation. Critical challenges facing international technology based firms.* Oxford University Press. New York.

Carlile, P. R., and Rebentisch, E. S. 2003. Into the black box: The knowledge transformation cycle. %*Management Science*,%49(9): 1180-1195.

Domingos, P., 2012. *A Few Useful Things to Know about Machine Learning*. Seattle: University of Washington.

Dugad, R. and Ahuja, N. 1998. *Unsupervised Multidimensional Hierarchial Clustering*. Illinois, Urbana, IL 61801: Beckman Institute, University of Illinois.

Enberg, C., Lindkvist, L., & Tell, F. 2006. Exploring the dynamics of knowledge integration: acting and interacting in project teams. %*Management Learning*,%37(2): 143-165.

Ernst, D., and Kim, L. 2002. Global production networks, knowledge diffusion, and local capability formation. *Research Policy*, 31(8-9): 1417-1429.

Feldman, F. 2011. What we learn from the experience machine. *The Cambridge Companion to Nozick's Anarchy, State, and Utopia*: 59-88.

Grant, R. M. 1996a. Toward a Knowledge-Based Theory of the Firm. *Strategic Management Journal*, 17 (Winter Special Issue): 109-122.

Grant, R. M. 1996b. Prospering in Dynamically Competitive Environments: Organizational Capability as Knowledge Integration. *Organization Science*, 7(4): 375-387.

Grant, R. M., & Baden Fuller, C. 2004. A knowledge accessing theory of strategic alliances. *Journal of Management Studies*, 41(1): 61-84.

Hall, P., Phan, W. and Whitson, K. 2016. *The Evolution of Analytics - Opportunities and Challanges for Machine learning in Business*. Sebastopol: O'Reilly Media Inc.

The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry *Daniel Viberg and Mohammad H. Eslami*

Hansen, M. T., Nohria, N., and Tierney, T. 1999. *What's your strategy for managing knowledge. The knowledge management yearbook 2000–2001.* 77(2): 106-116.

Hodson, J. 2016. How to make your company Machine Learning ready. *Harvard Business Review.*

Hoffer, E., and Ailon, N. 2015. Deep metric learning using triplet network. In *International Workshop on Similarity-Based Pattern Recognition.* Springer, Cham: 84-92.

Howells, J. (1996). Tacit knowledge. *Technology Analysis and Strategic Management*, 8(2): 91-106.

Huang, J. C., and Newell, S. 2003. Knowledge integration processes and dynamics within the context of cross-functional projects. *International Journal of Project Management*, 21(3): 167-176.

Jin, X., Wang, J., Chu, T., and Xia, J. 2018. Knowledge source strategy and enterprise innovation performance: dynamic analysis based on machine learning. *Technology Analysis & Strategic Management,*%30(1): 71-83.

Jonsson, A. (2012). *Kunskapsöverföring and Knowledge Managment.* Malmö: Liber AB.

Karypis, G., Han, E. H. S., and Kumar, V. 1999. Chameleon: Hierarchical clustering using dynamic modeling. *Computer*, 8: 68-75.

Kearns, M., Mansour, Y., and Ng, A. Y. 1998. An information-theoretic analysis of hard and soft assignment methods for clustering. In *Learning in Graphical Models*, Springer, Dordrecht: 495-520.

Kogut, B., and Zander, U. 1992. Knowledge of the firm, combinative capabilities, and the replication of technology. *Organization Science*, 3(3): 383-397.

Kubat, M. 2015. Artificial neural networks. In *An Introduction to Machine Learning.* Springer, Cham: 91-111.

Lin, B. W., and Chen, C. J. 2006. Fostering product innovation in industry networks: the mediating role of knowledge integration. *The International Journal of Human Resource Management*, 17(1): 155-173.

Li, J., and Herd, A. M. 2017. Shifting practices in digital workplace learning: An integrated approach to learning, knowledge management, and knowledge sharing. *Human Resource Development International*, 20(3): 185-193.

Liker, J. K., and Meier, D. 2006. *The Toyota Way Fieldbook; A Practical Guide for Implementing Toyota's 4P's.* Bok-McGraw-Hill Professional.

Marsh, S. J., and Stock, G. N. 2006. Creating dynamic capability: The role of intertemporal integration, knowledge retention, and interpretation. *Journal of Product Innovation Management*, 23(5): 422-436.

Miller, D. J., Fern, M. J., and Cardinal, L. B. 2007. The use of knowledge for technological innovation within diversified firms. *Academy of Management Journal*, 50(2): 307-325.

Nonaka, I. 1994. A dynamic theory of organizational knowledge creation. *Organization Science*, 5(1): 14-37.

Nonaka, I., and Takeuchi, H. 1995. *The knowledge-creating company: How Japanese companies create the dynamics of innovation.* Oxford University Press.

Okhuysen, G. A., and Eisenhardt, K. M. 2002. Integrating knowledge in groups: How formal interventions enable flexibility. *Organization Science*, 13(4): 370-386.

Provost, F. 2000. Machine learning from imbalanced data sets 101. In Proceedings of the *AAAI 2000 workshop on imbalanced data sets*, AAAI Press, 68: 1-3.

Schmickl, C., and Kieser, A. 2008. How much do specialists have to learn from each other when they jointly develop radical product innovations? %*Research Policy*, %37(3): 473-491.

Simon, H. A. 1983. Why should machines learn? *Machine Learning*. Morgan Kaufmann: 25-37.

Spender, J. C. 1996. Making knowledge the basis of a dynamic theory of the firm. *Strategic Management Journal*, 17(S2): 45-62.

Teece, D. J., Pisano, G., and Shuen, A. 1997. Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7): 509-533.

Tell, F. (2011). *Knowledge integration and innovation: a survey of the field. In Knowledge integration and innovation. Critical challenges facing international technology based firms.* Eds. Berggren C., Bergek A., Bengtsson L., Hobday M., So derlund J. Oxford University Press. New York.

Tiwana, A., and Mclean, E. R. 2005. Expertise integration and creativity in information systems development. *Journal of Management Information Systems*, 22(1): 13-43.

Yang, J. 2005. Knowledge integration and innovation: Securing new product advantage in high technology industry. %*The Journal of High Technology Management Research,*%16(1): 121-135.

Wiig, K. M. 1997. Knowledge management: where did it come from and where will it go? *Expert Systems with Applications*, 13(1): 1-14.

Witten, I. H., Frank, E., Hall, M. A., and Pal, C. J. 2016. *Data Mining: Practical Machine Learning tools and techniques.* Morgan Kaufmann.

The Effect of Machine Learning on Knowledge-Intensive R&D in the Technology Industry *Daniel Viberg and Mohammad H. Eslami*

About the Authors

Daniel Viberg has a M.Sc in Industrial Engineering and Management and a B.Sc in Mechanical Engineering from Linköping University in Sweden. He has experience in computer science from various spare time projects connected to both commercial and research purposes.

Mohammad H. Eslami is an Assistant professor at Jönköping International Business School in Sweden. His research interests are in the field of innovation management and knowledge integration. His research has been published in Industrial Marketing management, Journal of engineering and technology management, international journal of innovation management and etc.