Artificial intelligence and innovation management: A review, framework, and research agenda

Naomi Haefnera,⁎, Joakim Wincenta,b, Vinit Paridac,d,e, Oliver Gassmanna

a University of St. Gallen, St. Gallen, Switzerland
b Hanken, School of Economics, Entrepreneurship and Management, Helsinki, Finland
c Luleå University of Technology Luleå, Sweden
d University of South-Eastern Norway, Entrepreneurship and Innovation/USN Business School, Vestfold, Norway
e University of Vaasa, School of Management, Vaasa, Finland

ARTICLE INFO

Keywords:
Artificial intelligence  
AI  
Innovation  
Review  
Literature review  
Innovation management  
Machine learning  
Information processing

ABSTRACT

Artificial Intelligence (AI) reshapes companies and how innovation management is organized. Consistent with rapid technological development and the replacement of human organization, AI may indeed compel management to rethink a company’s entire innovation process. In response, we review and explore the implications for future innovation management. Using ideas from the Carnegie School and the behavioral theory of the firm, we review the implications for innovation management of AI technologies and machine learning-based AI systems. We outline a framework showing the extent to which AI can replace humans and explain what is important to consider in making the transformation to the digital organization of innovation. We conclude our study by exploring directions for future research.

1. Introduction

Scholarly interest in the idea that artificial intelligence (AI) and machine learning can replace humans, take over workplace roles, and reshape existing organizational processes has been growing steadily (Brynjolfsson and McAfee, 2017; von Krogh, 2018). The central premise is that, given certain constraints in information processing, AI can deliver higher quality, greater efficiency, and better outcomes than human experts (Agrawal et al., 2018a; Bughin et al., 2018).

Considering AI’s potential to take on traditional ‘human’ tasks in organizations, we may ask whether a role for AI can be used in pursuing one of the most important processes affecting a firm’s long-term survival and competitive advantage – innovation (Lengnick-Hall, 1992; Porter and Stern, 2001). Prima facie, the idea that AI and machine learning could and should be used by firms for innovation purposes may seem almost far-fetched. After all, innovation has traditionally been seen as a domain for humans, given their ‘unique’ ability to be innovative (Amabile, 2019).

Although AI may have downsides compared to humans, there are several non-trivial reasons why firms may want to use AI in their innovation processes. Among the factors exogenous to the innovation process is the fact that innovation managers are increasingly faced with highly volatile and changing environments, ever more competitive global markets, rival technologies, and dramatically changing political landscapes (Jones et al., 2016; O’Cass and Wetzel, 2018; Spieth et al., 2014). At the same time, the availability of information has increased and continues to increase significantly. These trends provide strong evidence that the baseline for competitiveness stands on the information and problem-solving capabilities of organizations (Hajli and Featherman, 2018). Perhaps more importantly, in many areas, the negative effects of innovation’s riskiness are being compounded by increasing costs. That is to say, the cost of each innovation has been increasing quite dramatically. For instance, while transistor density on integrated circuits has been increasing exponentially in line with Moore’s Law, this advance has necessitated ever greater efforts by firms such as Intel (Schilling, 2017). Drug development processes in the pharmaceutical industry show similar trends (Munos, 2009; Pammolli et al., 2011). This means that the way innovation is organized needs to be challenged by introducing AI and machine learning because of their cost advantages in information processing.

Consequently, finding ways to apply AI and machine learning to firms’ innovation processes should be of considerable interest to
innovation managers. On the one hand, this has the potential to create better ways for firms to respond to their increasingly competitive environment and manage the growing amounts of information around them. On the other hand, supporting the innovation process with AI could generate real value for firms by reducing both the riskiness and the costliness of innovation processes.

Today, human-organized innovation management plays a key role in companies and their capacity to reinvent themselves through exploratory initiatives. However, AI can provide instrumental assistance beyond the scope of humans (Groves et al., 2013; Wamba et al., 2017). Indeed, both academics and practitioners have asserted that AI may substantially impact firms’ innovation processes in the future (Buglin et al., 2018; von Krogh, 2018). The notion that AI could potentially be applied in innovation settings is further supported by the rapid development of AI and machine learning, which points to significant and intriguing changes to come (Lu, 2019; Varian, 2018; Ward et al., 2014). However, our knowledge of AI’s limitations in the context of innovation is still quite sparse. The use of AI and machine learning for creativity and innovation is very different from the established areas where AI has replaced traditional management (Chui et al., 2018).

Following on from the above discussion, the objective of this article is to fill the gap in our knowledge by reviewing the literature and offering a framework to examine management challenges associated with promoting innovation through AI. While AI has only recently started to gain momentum in the management literature (e.g., Adner et al., 2019; Bettis and Hu, 2018; Fan and Teodoridis, 2020; Goldfarb et al., 2020; Krakowski et al., 2019; Puranam et al., 2018; Raisch and Krakowski, Impress), the phenomenon is, of course, not new. When the idea of artificially intelligent computer systems was first discussed by experts in that field in the mid-1950s, the potential impact of computer processing on organizations was already of interest to management scholars, most notably Richard Cyert, James March, and Herbert Simon at the Carnegie School. In particular, the behavioral theory of the firm (BTF; Cyert and March 1963) has, since its inception, been intimately linked to AI (Augier and Prieto, 2007). Simon argued that “if computers are organized somewhat in the image of man, then the computer [is] an obvious device for exploring the consequences of alternative organizational assumptions for human behavior” (Simon, 1996, p. 21).

Our research offers a framework for explaining how AI can be used for innovative purposes, and it addresses calls to move beyond human involvement in the innovation process. In doing so, we build upon central assumptions of the behavioral theory of the firm and its key implications.

We proceed as follows. First, we provide the theoretical background to our study. We describe the link between the behavioral theory of the firm and artificial intelligence, paying special attention to organizational problem solving and information processing in this context. We also examine information processing in the digitized organization by elucidating the need for modern firms to compete on their digital capabilities and by explaining the new modalities of information processing in the digitized organization. In doing so, we describe the innovation process and the associated information processing constraints. Building on this theoretical background, we then examine potential AI application areas in the innovation process and derive a framework for overcoming information processing constraints in the innovation process with AI. We develop a set of readiness levels of AI in the digitized organization by looking at AI’s information processing capabilities. Then, we discuss our derived framework and the readiness levels by describing the different challenges in implementing AI in the innovation process. Finally, we draw some brief conclusions.

2. Theoretical background

The BTF has been acknowledged in organization theory and management as a major foundation for understanding decision making and organizational behavior (Argote and Greve, 2007). In developing it, Cyert and March (1963) proposed a set of foundational concepts on the cognitive level that are built on the concept of bounded rationality, which encapsulates the ideas of satisfying, search, and organizational routines. The theory includes a set of relational concepts that serve as theoretical mechanisms to explain how cognitive concepts unfold in organizations. These concepts include the quasi-resolution of conflict, uncertainty avoidance, problemistic search, and organizational learning (Gavetti et al., 2012). There is renewed interest among researchers in re-examining the various concepts put forward by Cyert and March in ‘A Behavioral Theory of the Firm’ (Piezunka et al., 2019; Posen et al., 2018; Puranam et al., 2015) in the context of recent developments in AI.

The idea originally posited by the BTF is that organizational problem solving could be better understood by looking at organizations as information-processing systems constructed by simple computational ‘if-then’ algorithms, which were at the core of AI at that time. The logic of viewing the organization as a simple algorithm or a combination of algorithms that process information is deeply embedded in the BTF (Cyert and March 1963).

2.1. The behavioral theory of the firm and information processing

Information processing is a key component in innovation in organizations. A central activity in innovation management is the process of decision making, which requires information processing by managers involved in the innovation process (McNally and Schmidt, 2011; van Riel et al., 2004). The role of management in information processing is to decide upon inputs into the process in terms of data, knowledge, and other information. Then, information must be processed – in other words, data, knowledge, and information are gathered and analyzed. Finally, once information has been processed, management has the responsibility to take decisions.

With the advent of machine learning – a type of AI that allows machines to ‘learn’ from data and experience without being explicitly programmed (Samuel, 1959) – the way information processing occurs in organizations is changing rapidly. All the above stages of organizational information processing can be supported or, in some cases, taken over by AI systems. Indeed, the modern digitized organization exhibits certain characteristics that substantially change the way information processing occurs in organizations. Interestingly, the organizations of today are changing in a way that makes it difficult for management to obtain and analyze certain elements of information.

2.2. Information processing in the digitized organization

The digitized organization that has now emerged features a strong backbone of highly integrated machine learning and computerized knowledge. This means that a vast number of processes are automated through algorithms. Some authors suggest that this needs to be an organizational mainstay and, therefore, organizations should consider their core capabilities as digital capabilities (Lenka et al., 2017). These services interact with customers and suppliers, and enable the storage of information and knowledge (George et al., 2014; Lanzolla et al., 2018; Zammuto et al., 2007). Thus, an increased amount of information and knowledge is stored electronically and without human involvement. The digitized organization becomes the major constituent, and the social system of an organization becomes less pivotal. Consequently, one can say that executives and directors who are responsible for innovation management and decision making are less efficient not only because of human limitations but also because they may be constrained by operating outside the relevant flow of information. It can be assumed that those managers who do have access to this information are a small subset of the managerial pool, which means that many managers may have quantitatively and qualitatively less information than they had prior to the computerized organization and the
technological changes in the workplace.

These background realities call for a model where innovation-oriented AI and machine learning of computerized information and processes are integrated into innovation management. As AI advances further, it can be said that the role of innovation management will change in step with progress made by AI and machine learning. Thus, human innovation management will be expected to work side by side with AI and machine learning algorithms in identifying and selecting opportunities as well as investigating what could be the organization’s next competitive advantage.

2.3. Information processing in the innovation process

To better understand how AI augments organizational innovation, we need to examine how information is processed for innovation. The innovation process—which is at the core of innovation management's attention—is commonly understood to comprise a series of stages including (1) the recognition, discovery, creation, and generation of innovative ideas, opportunities, and solutions; (2) the development or exploitation of various ideas, opportunities, and solutions; and finally (3) the evaluation and selection of one or several of the most promising ideas, opportunities, and solutions (e.g., Kijkuit and van den Ende, 2007). It can be argued that the first two steps in particular require significant levels of creativity and out-of-the-box thinking (Martin and Wilson, 2016; Shane, 2000). Since we are interested in determining where and how AI can be used to support human decision making in the innovation process, we will focus on the first two stages of the process—namely, idea generation and idea development.

We believe that the increased implementation of electronic services and automation coupled with the general transformation to digitized organizations will change the role of innovation management. As in the past, when innovation managers attempt to recognize or develop new opportunities and ideas, they face two specific barriers (Eggers and Kaplan, 2009). First, they must overcome information processing constraints (Nelson and Winter, 1982; Williams and Mitchell, 2004) that limit the amount of information on either new opportunities or possible solutions the firm may pursue. These information processing constraints are often the result of managers' cognitive limitations—that is to say, human mental capacities to absorb or process information are biologically limited. The second barrier encountered by managers is the result of ineffective or local search routines (Gavetti and Levinthal, 2000; Katila and Ahuja, 2002). This barrier specifies that managers generally search for solutions in knowledge domains that are related to the firm's and their own existing knowledge base (Posen et al., 2018). This suggests that most solutions will be comparatively incremental in their innovative thrust since they rely very closely on existing knowledge. However, to generate a more creative and innovative idea or opportunity, managers will have to extend search beyond existing knowledge domains to new fields that are more exploratory in nature.

Therefore, even though access may be more limited in increasingly digitized organizations, the more managers are able to process a large amount of information on possible solution approaches and opportunities, the more they should be able to whittle down the set of possible solutions to the most promising ones and to recognize truly exciting opportunities. Furthermore, since managers are able to go beyond their current knowledge base with the assistance of AI, they should be able to develop more innovative solutions and recognize more creative opportunities (Amabile, 2019; von Krogh, 2018). The AI solutions that could be employed are not straightforward however, and it may be challenging to involve AI in the innovation process. It will also be difficult to replace human involvement. Any artificial intelligence-based system that seeks to support management in these endeavors must be capable of overcoming the same barriers encountered by human managers in the innovation process.

The above discussion develops the fundamental perspective used for a framework to examine management challenges associated with promoting innovation through AI. Table 1 below provides an overview of the literature streams and topics covered in our theoretical background section. We bring together the behavioral theory of the firm and its focus on information processing with the literature on digitized organization and innovation processes to theorize about the challenges that management faces with AI and innovation. Next, we turn to the specific analysis.

3. Potential AI application areas in the innovation process

By combining the barriers that must be overcome by both humans and AI systems in the innovation process with the key activities of idea generation and development that need to be conducted, we can derive a framework of potentially creative application areas of AI within the innovation process. To understand the possibilities of AI, we need to delineate where AI can assist and potentially replace human decision making in innovation management. Specifically, there are four potential areas where human decision making could theoretically be supported: (1) developing ideas by overcoming information processing constraints; (2) generating ideas by overcoming information processing constraints; (3) developing ideas by overcoming local search routines; and (4) generating ideas by overcoming local search routines. These four areas are depicted in Fig. 1 along with a brief description of what AI in each quadrant ought to be capable of doing.

The next section provides an overview of the current capabilities of AI systems in supporting humans in the aforementioned areas of the innovation process by delineating examples in each quadrant of Fig. 1.

3.1. Overcoming information processing constraints with AI to develop ideas

Current AI systems excel at overcoming humans' information processing constraints in the area of idea and opportunity development. Currently, AI systems rely heavily on deep neural networks that require, and are able to process, vast amounts of data (Ng, 2017). With this feature, we see a veritable plethora of AI systems that are able to support humans in the development of ideas, opportunities, and solution approaches by processing a much larger amount of information than is humanly possible and unearthing interesting areas for investigation. Indeed, these technologies are already creating substantial economic value for firms (Roose, 2019). In this area, referring to quadrant 1 in Fig. 1, we find a number of interesting applications of AI across a very wide range of domains. This development is strongly linked to improved conditions for innovation. There are many exciting applications of AI systems in materials discovery. For instance, AI can be used to optimize battery components and solar cells (Charlington, 2018), or to speed up the discovery process for new catalysts (Tran and Ulissi, 2018). In order to discover these new materials, machine learning-based methods are used to predict the most promising materials to test, thereby speeding up the innovation process substantially. There are, of course, interesting AI applications in pharmaceutical research and development as well (e.g., Mamoshina et al., 2016; Schuhmacher et al., Inpress). Here, AI systems include uses that speed up the process of protein engineering (Yang et al., 2019), which is instrumental in discovering proteins suitable for technological, scientific, and medical applications. The reason why methods based on machine learning are interesting to researchers in this domain is because the search space of possible proteins is too large to search exhaustively with existing methods (Yang et al., 2019). Furthermore, AI...
applications can be used to identify treatments for disease – for example, deep domain adaptation neural networks have been trained on single cell RNA genomics datasets to ultimately develop treatments that will stop the transmission of malaria (Johansen and Quon, 2018). Finally, there are many areas where AI systems can be used to create process innovations in organizations. For instance, Celonis uses process mining to identify organizational processes that are suitable for robotic process automation (Geyer-Klingeberg et al., 2018; Veit et al., 2017). Thus, Celonis uses AI applications that enable organizations to implement significant administrative innovations.

3.2. Overcoming information processing constraints with AI to generate ideas

There are several AI applications that relate to quadrant 2 of the framework in Fig. 1. These AI applications are able to process much more information to generate new ideas and opportunities that would likely be overlooked by humans operating on their own. A typical example is an application developed by Outlier.ai. The company uses a suite of machine learning methods to process raw metrics data into insights that are humanly readable (Unemyr, 2018). After analyzing a firm’s data, Outlier generates a set of customized ‘stories’ that summarize actionable and interesting insights for specific managers. In doing so, Outlier can highlight innovative opportunities for managers. How this can work is illustrated in the following example:

“One of Outlier’s customers is a large, international quick-service restaurant franchise. It sells hundreds of items across thousands of stores but, in one instance, the company found something that was different. A store that had closed for three weeks was immediately selling twice as many fountain drinks as before when it reopened. This is a big change since fountain drinks are the highest-margin items sold by quick-service restaurants. Upon further investigation, the management found this one location had been closed for renovations but after reopening had not been reset to the previous layout. The staff had found a better layout for the store, one that drove significantly more fountain drink sales, by accident. This single observation from a single store, of a change made by accident, can now change the entire revenue of the company as it’s rolled out across all locations.” (Byrnes, 2018)

As is evident from this example, the AI-based analysis provided by Outlier was instrumental in developing an innovation in the focal firm. Outlier’s ability to find anomalies and significant patterns in business data is one way in which AI can assist firms in generating or recognizing innovative ideas and opportunities. These AI methods may not be able to independently develop entire solutions, but they can point human managers towards the most promising avenues for innovation.

Another interesting example in this area is provided by Tshitoyan and colleagues (2019). They created an AI system that can capture latent knowledge from the materials science literature. Their system uses the word2vec algorithm – a popular neural network in natural language processing applications – to derive embeddings of concepts in the literature. The algorithm is able to capture complex materials science concepts – including the underlying structure of the periodic table – without any explicit insertion of chemical knowledge by the researchers. The AI system can also recommend materials for functional

| Literature stream | Topic | Authors |
|-------------------|-------|---------|
| Behavioral theory of the firm (BTF) | General overview | Argote and Greve (2007), Cyert and March (1963), and Gavetti et al. (2012), Piezunka et al. (2019), Posen et al. (2018), and Puranam et al. (2015). |
| Information processing | Importance for innovation in organizations | McNally and Schmidt (2011) and van Riel et al. (2004). |
| Digitized organizations | Machine learning capabilities | Samuel (1999). |
| Innovation process | Digital capabilities | Lenka et al. (2017). |
| | New modalities of knowledge and information management | George et al. (2014), Lanzolla et al. (2018), and Zammuto et al. (2007). |
| | Steps and characteristics | Kijkuit and van den Endes (2007); Martin and Wilson (2016); Shane (2003), Elgers and Kaplan (2009), Nelson and Winter (1982), Williams and Mitchell (2004), Gavetti and Levinthal (2000), Katila and Ahuja (2002), and Posen et al. (2018), Amabile (2019) and von Krogh (2018). |

Table 1
Overview of literature streams and topics.

| Literature stream | Topic | Authors |
|-------------------|-------|---------|
| Information processing constraints | AI system is able to identify and evaluate more information that can then be used to develop ideas. | (1) |
| Ineffective or local search routines | AI system is able to identify and evaluate more creative/exploratory ideas. | (3) |
| | AI system is able to recognize and create more creative/exploratory problems, opportunities, and threats to generate new ideas. | (4) |
| | AI system is able to recognize more problems, opportunities, and threats that may be used to generate new ideas. | (2) |

Fig. 1. Application areas of AI in the innovation process.
applications. By censoring the data, the authors can show that the system is, in fact, able to recommend materials several years before their discovery. Thus, this method points to potential opportunities for future innovations, albeit within an already existing knowledge domain. This study is indicative of potential AI applications in quadrant 2 of Fig. 1 – namely, AI systems that are able to assist in generating or recognizing ideas and opportunities for innovation where a large amount of information in an existing domain of knowledge has to be processed.

3.3. Overcoming local search routines with AI to develop ideas

There is some initial evidence that AI systems may be able to support humans in the types of innovative activity represented in quadrant 3 of Fig. 1. These activities entail identifying and developing ideas, opportunities, and solution approaches where the process goes beyond using local search routines – in other words, distant search is used. Autodesk, for instance, used various algorithms to create a new crew partition for Airbus (Autodesk, 2016). The generative design methods employed to devise the new partition create the kinds of product that designers could not conjure up on their own (Rhodes, 2015). The algorithms used by Autodesk were based on the growth patterns of slime mold and mammal bones. They enabled the construction of a new, more efficient, but equally stable crew partition. Thus, by incorporating AI methods into the development process, Autodesk and Airbus were able to generate a more innovative solution than would have been otherwise possible.

Even more interesting are some applications based on generative adversarial networks (GANs). The creative adversarial network (CAN) for art creation developed by Elgammal and colleagues (2017) is an example of such an AI solution. The CAN is a type of GAN that is able to generate novel art. The network is trained on 81,449 paintings from 1119 artists ranging from the 15th to the 20th century. The system trains two competing networks – a discriminator and a generator – to learn art style classification (discriminator) and style ambiguity (generator). As a result, the CAN generates new art that deviates from the learned styles. We would argue that this deviation from previously learned styles is precisely where the CAN system is able to overcome local search routines and show its potential for distant search.

There is some initial evidence that AI systems may be able to support humans in the types of innovative activity represented in quadrant 3 of Fig. 1. These activities entail identifying and developing ideas, opportunities, and solution approaches where the process goes beyond using local search routines – in other words, distant search is used. Autodesk, for instance, used various algorithms to create a new crew partition for Airbus (Autodesk, 2016). The generative design methods employed to devise the new partition create the kinds of product that designers could not conjure up on their own (Rhodes, 2015). The algorithms used by Autodesk were based on the growth patterns of slime mold and mammal bones. They enabled the construction of a new, more efficient, but equally stable crew partition. Thus, by incorporating AI methods into the development process, Autodesk and Airbus were able to generate a more innovative solution than would have been otherwise possible.

Even more interesting are some applications based on generative adversarial networks (GANs). The creative adversarial network (CAN) for art creation developed by Elgammal and colleagues (2017) is an example of such an AI solution. The CAN is a type of GAN that is able to generate novel art. The network is trained on 81,449 paintings from 1119 artists ranging from the 15th to the 20th century. The system trains two competing networks – a discriminator and a generator – to learn art style classification (discriminator) and style ambiguity (generator). As a result, the CAN generates new art that deviates from the learned styles. We would argue that this deviation from previously learned styles is precisely where the CAN system is able to overcome local search routines and show its potential for distant search. Since the model initially learns about existing art styles, it is knowledgeable about current domain knowledge. However, it is set up to specifically explore beyond current styles and is, therefore, able to generate novel ideas. Another related research project by Sbai and colleagues is called DesIGN – design inspiration from generative networks (2018). This system can generate novel styles, forms, and shapes for fashion apparel. DesIGN deviates from existing fashion styles as represented in the training dataset, while generating realistic pieces of clothing. It, therefore, overcomes local search routines when developing new ideas for fashion apparel.

3.4. Overcoming local search routines with AI to generate ideas

Finally, AI systems hoping to address quadrant 4 of Fig. 1 must be able to generate or recognize ideas and opportunities for innovation in unrelated knowledge domains. A method in artificial intelligence that may facilitate the generation or recognition of innovative ideas and opportunities is reinforcement learning. There have been recent advances in reinforcement learning such as unsupervised reinforcement learning and meta-reinforcement learning that could conceivably be helpful in generating novel ideas. Reinforcement learning in general involves training an agent in a (virtual) environment. The agent uses a reward signal to learn which actions maximize rewards and which actions diminish them. Reinforcement learning requires humans to handcraft rewards, which is a non-trivial and sometimes sub-optimal approach to reward engineering. As Simon Osindero, a top AI researcher from Google DeepMind, explains, “To the extent that you're hand designing a reward function, you're also in some sense hand designing a solution [...] If it was easy for us to design a solution, then maybe you wouldn't need to learn it in the first place” (Charrington, 2019). Unsupervised reinforcement learning tries to address this shortcoming by allowing the agent to learn its reward function using a stream of observations and actions (Warde-Farley et al., 2018). Thus, this method is a first step toward enabling algorithms to learn to recognize and achieve goals without any supervision, which will open up interesting avenues for creativity and innovation. Meta-reinforcement learning tackles a closely related question concerning how learning can be used to improve the process of learning itself. Recent work in this area (e.g., Gupta et al., 2018) has attempted to devise algorithms that are able to adapt rapidly to arbitrary new problems (Charrington, 2019). Advancements in these areas should allow algorithms to become more flexible in terms of solving new problems, which may prove helpful in generating, discovering, and recognizing new creative ideas and opportunities.

4. AI readiness levels for developing the digitized organization

As foreshadowed above, the different AI systems described in section 3 are at different levels of sophistication in terms of their ability to augment and replace human managers in innovation processes. These levels of sophistication can be derived by looking at the kinds of capabilities that an information processing system must have in order to complete the functions described in each of the quadrants in Fig. 1. For this, we will consider the ‘innovation process’ and the ‘barriers to innovation’ dimensions as the problem space and the solution space, respectively (Gotta and Troya, 1998; Restrepo and Christiansen, 2004).

The first dimension, which describes the tasks in the innovation process (idea development and idea generation), can also be viewed as the problem space that is the subject of innovation. In line with an information processing perspective on the innovation process, the “problem space is the internal representation of the task environment” used by the subject (Newell and Simon, 1972, p. 56), the subject being a human manager or an AI system. When going through the innovation process, an information processing system can either continue with its current definition of the problem space, which would correspond to simply developing a new idea or solution based on the problem space, or it could decide to include additional data, information, and/or knowledge, thereby redefining the problem space and opening up the ability to generate new ideas and solutions (Newell and Simon, 1972, p. 88). Another way to describe these two options would be to consider the former as the exploitation of an existing problem space and the latter as the exploration of a redefined, evolving, or different problem space.

The second dimension, describing the barriers to overcome in the innovation process (information processing constraints and ineffective or local search), may be interpreted as the ways in which the solution space for innovation can be altered. Overcoming information processing constraints does not require any change in the specification of the solution space, since this barrier to innovation ‘merely’ indicates that the solution space is searched more efficiently and quickly. In other words, overcoming information processing constraints indicates that the solution space is more effectively and efficiently exploited. In order to overcome local and inefficient search routines, however, it is necessary to explore the solution space so that more distant and creative solutions can be found.

To understand the capability levels of current AI systems in terms of assisting humans in the innovation process, it is important to understand some key technical features of these systems. Specifically, there are two key characteristics in most AI systems developed today that are constrained by human capabilities. First, most current AI systems are trained by human AI experts who partner with domain experts relying on their existing knowledge base. This means that these AI systems should generally try to search a known, related knowledge base more extensively – that is to say, most systems are limited in the extent to
which they can explore the problem space. Second, state of the art AI systems are set up so that the learning process is optimized for a given objective function (Goodfellow et al., 2016). This objective function is defined by the human AI researchers implementing and training the system. Moreover, these objective functions are generally very sparse since the human researchers who are calibrating the systems cannot possibly know all possible objectives and, therefore, tend to fall short in their ability to provide an ideal objective function. Consequently, for most AI applications, the solution space is pre-defined by humans, and so current AI systems tend to have a very limited ability to explore the solution space autonomously.

As a result, these two features of AI systems pose technical limitations on the systems’ abilities to redefine and explore both the problem space and the solution space. Furthermore, most current AI systems are limited in their ability to generate or recognize ideas and opportunities and to overcome local search routines. However, as explained in sections 3.2, 3.3, and especially 3.4, there have been some recent advancements suggesting that AI systems may indeed be able to overcome these limitations. Thus, we are able to derive a range of what we term ‘information processing capability levels’ for AI systems that indicate how likely AI systems are to replace and complement human decision making. Broadly, they can be grouped into three capability levels according to the types of information processing capabilities they display, as depicted in Fig. 2 below.

4.1. Information processing capability level 1: Exploiting

Information processing capability level 1 indicates that the AI system is capable of helping human innovation managers to process much larger amounts of information and knowledge than they would be able to accomplish on their own. AI systems at this capability level will primarily be able to support rather than fully replace humans in the innovation process because, by processing more information, they are exercising a supporting function and not fully taking over the entire innovation process. These AI systems are able, therefore, to help humans to overcome the cognitive information processing limits that often hinder them from fully considering vast amounts of data and paying attention to a multitude of data sources. Properly designed AI systems can both deal with much larger amounts of data and process many different data sources. These types of AI system are located in quadrant 1 of the Fig. 1 framework presented in section 3.

4.2. Information processing capability level 2: Expanding

Information processing capability level 2 implies that the AI system is capable of either expanding the innovation process by generating new ideas and opportunities or by overcoming local search routines to find more distant solutions. At this capability level, AI systems are still working in tandem with human innovation managers. These systems excel at supporting managers in two particular ways. First, they help in discovering new ideas and opportunities as described in quadrant 2 of the Fig. 1 framework. Second, they can support innovation managers in developing more innovative and creative ideas and solutions. These types of AI system are depicted in quadrant 3 of Fig. 1. At the moment, the technological capabilities of AI systems are still relatively limited, and only a few systems are able to actually function at this readiness level, as explained in sections 3.2 and 3.3.

4.3. Information processing capability level 3: Exploring

Information processing capability level 3 signals that the AI system is capable of exploring new avenues in the innovation process. These types of AI system can accomplish more advanced and difficult tasks in the innovation process and, therefore, are able not only to support human innovation managers but also to replace them to a certain extent. AI systems at the ‘exploring’ information processing capability level can generate and create new ideas that are especially innovative and creative. Due to their more advanced information processing capabilities, these AI systems are able to explore both new ways of defining problems (exploring the problem space) and new ways of addressing the problem (exploring the solution space). Consequently, we can expect AI systems with information processing capability level 3 to have a greater chance of being able to take over a larger share of the tasks traditionally undertaken by human innovation managers. However, the current state of the art is relatively far removed from allowing the implementation of such AI systems because there are few initial forays into AI systems of this kind. This is explained in section 3.4 with regard to quadrant 4 in the Fig. 1 framework.

5. Discussion

Considering the opportunities to involve AI in the innovation process, the question of when, how, and to what extent human innovation managers and AI systems can and should work together arises. This has been discussed in the literature but usually from the perspective of simply understanding AI’s ability to perform and replace human workplace tasks in general. For instance, current analyses estimate that proven AI technologies have the potential to replace up to half of all work activities carried out by humans – 60 percent of all occupations consist of approximately 30 percent automatable activities (Bugain et al., 2017). Consequently, we think it important to undertake a more specific discussion of AI’s ability to replace humans in the innovation process.

Could AI ever replace the human side of innovation management? An initial investment in AI will generate fast, inexpensive, and relatively thorough manifestations of new ideas that can be innovative. However, the judgement of managers may be difficult to replace and, therefore, a full transformation to a digitized organization may be problematic. Developing and adding new innovations is often coordinated by a large management team that is motivated to explore market opportunities. In this regard, we must stress that innovation management decisions throughout the organization are inherently complex and, therefore, difficult to fully replace by AI. It would require a host of algorithms to be interwoven and, inescapably, this would be done under conditions of significant uncertainty. This is an art that would require the company to exploit economies of scope (Teceo, 1980), increase and build market power (Caves, 1981), and create flexible shifts and synergies with resources such as labor throughout the business areas of the company (Hill and Hoskisson, 1987). The full use of AI is challenging because it demands new ways of addressing a novel industry environment (Ransbotham et al., 2017; Yun et al., 2019). It entails the acquisition of new knowledge and resources, as well as creating new business logics and new business models to meld the new innovations into the current portfolio of products (Brynjolfsson and McAfee, 2017; Prahalad and Bettis, 1986). Companies would need to create and adjust routines that align with the new product, configure new organizational structures and systems for the purposes of administrative alignment, and buttress governance control. These are all tasks and activities that can be supported by AI but within clear and challenging boundaries.

While AI may help with the product concept and market analysis, and the scheduling of resources and the systems around it, it is a highly complex process. Thus, AI is likely to be more relevant when new products are launched in areas where the top management team (TMT) is less familiar. However, its use is likely to run alongside human management. Previous research has reported that overburdened and stressed management may not be able to develop sufficient knowledge to become familiar with new products, taking ill-informed decisions that are difficult to revise and that ultimately spell failure (Gary, 2005). The use of AI will likely make an important contribution to profitability when radically innovative products are launched and when the role of the TMT is different in the future.
How provisional are AI solutions and how difficult are they to implement? There are several challenges associated with implementing these emerging technologies in organizations. The specific challenges are located on the level of the technology itself as well as on the level of the individuals tasked with implementing it. Certain challenges are also located at the technology–human nexus.

The first set of challenges, which are closely related to the technology itself, include some rather more obvious challenges such as the issue of data availability and suitability (for an extensive discussion of this, see Agrawal et al., 2018b). On the technological side, there is the issue of hardware. For instance, in terms of compute power, some modern AI applications require extremely powerful processing functionalities and vast amounts of data to power these processes (CB Insights, 2019). For instance, one recent research project that generated fake images using generative adversarial models required as much energy as the average American household would use in approximately six months (Schwab, 2018). Beyond these challenges, the technology is in many ways not mature enough to be applied to professional settings. Taking reinforcement learning as one example, this area of machine learning is highly vibrant, and researchers are continuing to make very interesting progress (Charrington, 2019). However, while reinforcement learning is a highly researched and interesting area of AI, it is mostly applied to the development of AI systems that can beat human performance in video games. To date, there are only a few commercial applications of this very interesting type of AI. One example of a real-world application of reinforcement learning is its use by DiDi Chuxing, China’s biggest ride-hailing company (Lin et al., 2018; Qin et al., 2019). Didi has developed a reinforcement learning-based dispatching

| SEARCH APPROACH | LEVEL 1: EXPLOITING | LEVEL 2: EXPANDING | LEVEL 3: EXPLORING |
|-----------------|---------------------|---------------------|---------------------|
|                 | Able to successfully exploit both the problem and the solution spaces | Able to explore and redefine either the problem or the solution space | Fully able to explore and redefine both the problem and the solution spaces |

| CHARACTERISTICS | - Used to overcome cognitive information processing constraints  
- Can deal with more data  
- Able to process many different data sources | Either:  
- Able to discover new ideas and opportunities  
Or:  
- Supporting humans in developing more innovative ideas and solutions | - Exploring new avenues in the innovation process  
- Generate and create innovative and creative new ideas  
- Explore new ways of defining problems  
- Explore new ways of addressing problems |

| MATURITY LEVEL | Realizable applications | Initial implementations | Sandbox experiments |

| AUTONOMY LEVEL | Human-designed AI systems | Al systems with increasing machine autonomy |

Fig. 2. Information processing capability levels of artificial intelligence.
algorithm that is able to adapt to rider demand. The solution has been
tested in a limited number of Chinese cities where it showed greater
efficiency than prior non-reinforcement learning-based dispatching
systems (Hao, 2018). Aside from the fact that many machine learning
applications have not progressed substantially beyond sandbox envi-
nvironments, the technology itself is still undergoing development of its
fundamentals. Deep Learning was arguably proven to be viable only in
2012 (Parloff, 2016), and a large portion of the patents in AI are still
very much foundational in nature (EconSight, 2019).

The second set of challenges are closely related to the humans in-
volved in implementing and using AI solutions in firms. It is quite well
documented that firms often lack the necessary technical skills to suc-
scessfully implement AI solutions (Chui and Malhotra, 2018). Depending
on the complexity of the solution to be developed, different skills are
necessary and, since there is very high demand for these skills, com-
panies often have trouble acquiring the necessary talent. Companies
that do have employees with the necessary technical skills then en-
counter the next hurdle. If high-performing AI solutions are to be de-
veloped, the team working on the solution should generally comprise
both technical employees and domain experts (Bughin et al., 2018;
Daugherty and Wilson, 2018). The problem is that such collaborative
approaches to developing AI solutions can be quite complex. A recent
multi-year project to monitor patients in intensive care units (ICUs)
necessitated close collaboration between AI researchers and medical
professionals. This meant the amount of time required and the com-
plexity level of conducting the study were much higher than traditional
AI projects. But this approach was critical in order to design an effective
system (Yeung et al., 2019). Collaborative teams such as the one em-
ployed in the ICU monitoring project are essential to ensure that the AI
solutions developed address relevant problems that firms are currently
facing.

Finally, there are some challenges located at the nexus of the technol-
ogy and the humans in charge of implementing it. For example, a
limiting factor in applying AI systems in firms may stem from the
amount of human intervention required. While AI solutions are in-
tended to automate processes in workflows, it is seldom the case that a
whole series of connected tasks can be fully automated. Furthermore,
the solution space that AI systems can explore is, in many cases, very
much pre-defined by the algorithm(s) chosen by the humans im-
plementing the system. In addition to limiting the solution space, hu-
mans can also underspecify solutions. This is often the case with the
sandbox applications of reinforcement learning where sparse reward
functions lead to very ‘creative’ problem solving by the algorithm – the
machine essentially ends up gaming the system. Inadequate specifi-
cations by humans can also lead to questionable results in generative
design. When parameters are not stringent enough, the results can be so
‘creative’ as to be largely useless. Consequently, human intervention is
required but that has the potential to spawn inefficiencies in the pro-
cesses. Nonetheless, human intervention can be beneficial depending
on the context. One of the biggest challenges is, therefore, gaining a
clear understanding of when to circumvent human intervention and
when to embrace it. Furthermore, it is important to ensure that humans
receive actionable information from the AI system so that they can make
optimal decisions based on machine output. Another challenge
located at the human–technology nexus is that of trust in the AI system.
Depending on the design of the AI system, humans can sometimes trust
the technology either too much or too little, which creates friction in
using the AI system (Gilkinson and Woolley, 2020). Therefore, designing
AI systems that humans who interact with them can adequately trust is
an important challenge to overcome when implementing AI systems.

6. Conclusion

In this article, we review how innovation management may be
supported by artificial intelligence systems. Human-centered, conven-
tional approaches to innovation management have limitations that are
primarily rooted in their imperfect ability to fully address information
needs and cope with complexity. We developed a framework based on
information processing constraints as presented in the behavioral
theory of the firm. From that, we then derived the information pro-
cessing capability levels of AI needed to develop digitized organiza-
tions. Finally, we delineated the challenges in implementing AI systems
that innovation management faces in relation to the technology itself,
the humans tasked with implementing it, and the technology–human
nexus. Overall, we note that AI has a constructive role to play where the
tried-and-true benefits of innovation management resources are over-
whelmed, are impossible because of digitization, or when AI emerges
irrefutably as the preferred option. From our observations, it appears
that the clear potential of AI resides in creating a more systematic ap-
proach by integrating AI into organizations that are pursuing innova-
tion. Our research advances the innovation management literature by
shedding light on the use of AI and machine learning algorithms in the
future organization of innovation. Our findings point to areas where AI
systems can already be fruitfully applied in organizational innovation –
namely, instances where the development of new innovations is pri-
marily hampered by information processing constraints. AI systems that
rely on anomaly detection, for instance, can be helpful when firms are
struggling with information processing constraints as they search for new
opportunities. Finally, we highlight recent advancements in AI
algorithms that are indicative of AI’s potential to resolve the more
difficult challenges in innovation management. These include over-
coming local search and generating completely novel ideas. We look
forward with interest to see how new developments in AI technology
open up further possibilities and extend the areas where AI can usefully
be applied in innovation management.

CRediT authorship contribution statement

Naomi Haefner: Conceptualization, Investigation, Writing - origi-
inal draft, Writing - review & editing. Visualization, Project adminis-
tration. Joakim Wincent: Conceptualization, Writing - original draft.
Writing - review & editing, Supervision. Vinit Parida: Writing - original
draft, Writing - review & editing. Oliver Gassmann: Supervision,
Writing - review & editing.

References

Adner, R., Puranam, P., Zhu, F., 2019. What is different about digital strategy? From
quantitative to qualitative change. Strateg. Sci. 1–22. https://doi.org/10.1287/stsc.2019.0099.
Agrawal, A., Gans, J., Goldfarb, A., 2018a. Exploring the impact of artificial intelligence:
prediction versus judgment (No. 24629). In: NBER Working Paper, NBER Working
Paper. Cambridge, MA. https://doi.org/10.3386/w24629.
Agrawal, A., Gans, J., Goldfarb, A., 2018b. Prediction Machines: The Simple Economics of
Artificial Intelligence. Harvard Business Review Press, Cambridge, MA.
Amabile, T., 2019. Creativity, artificial intelligence, and a world of surprises. Acad.
Manag. Discov. 2019.0075. https://doi.org/10.5465/aml.2019.0075.
Argote, L., Greve, H.R., 2007. A Behavioral Theory of the Firm —40 years and counting:
introduction and impact. Organ. Sci. 18, 537–549. https://doi.org/10.1287/orsc.1070.0280.
Augier, M., Prietula, M., 2007. Historical roots of the “A Behavioral Theory of the Firm”
model at GSIA. Organ. Sci. 18, 507–522. https://doi.org/10.1287/orsc.1070.0276.
Autodesk, 2016. Airbus: reimagining the future of air travel [WWW Document]. Autodesk
URL: http://www.autodesk.com/customer-stories/airbus (accessed 1.29.19).
Bettis, R.A., Hu, S., 2018. Bounded rationality, heuristics, computational complexity,
and artificial intelligence. Adv. Strateg. Manag. 39, 139–150. https://doi.org/10.1108/
10742322018890093010.
Brynjolfsson, E., McAfee, A., 2017. The business of artificial intelligence: what it can –
and cannot – do for your organization. Harv. Bus. Rev. Jul 10.
Bughin, J., Hazan, E., Ramaswamy, S., Chui, M., Atlas, T., Dahstrom, P., Henke, N.,
Trench, M., 2017. Artificial intelligence: The next digital frontier? McKinsey Global
Inst. https://www.mckinsey.com/~/media/mckinsey/industries/advanced
%20electronics/our%20insights/how%20artificial%20intelligence%20can%20
deliver%20real%20value%20to%20companies/mgi-artificial-intelligence-
discussion-paper.axd.
Byrnes, S., 2018. The importance of having an AI-powered business strategy. Forbes.
https://www.forbes.com/sites/scottbyrnes/2018/05/02/advances-but-foundational-barriers-
remains. https://www.cbinsights.com/insights/artificial-intelligence-ai-adoption-
advances-but-foundational-barriers-remains.
CB Insights, 2019. Artificial Intelligence Trends. CB Insights New York. https://www.
