Journal of Product Innovation Management / Volume 40, Issue 2 / p. 139-153

CATALYST | Open Access

Augmenting human innovation teams with artificial intelligence: Exploring transformer-based language models

Sebastian G. Bouschery, Vera Blazevic, Frank T. Piller

First published: 10 January 2023
https://doi-org.ezproxy.auckland.ac.nz/10.1111/jpim.12656

Funding information: German Federal Ministry of Economic Affairs and Climate Action (BMWK), Grant/Award Number: 01MK20015A; Deutsche Forschungsgemeinschaft, Grant/Award Number: 390621612

Abstract

The use of transformer-based language models in artificial intelligence (AI) has increased adoption in various industries and led to significant productivity advancements in business operations. This article explores how these models can be used to augment human innovation teams in the new product development process, allowing for larger problem and solution spaces to be explored and ultimately leading to higher innovation performance. The article proposes the use of the AI-augmented double diamond framework to structure the exploration of how these models can assist in new product development (NPD) tasks, such as text summarization, sentiment analysis, and idea generation. It also discusses the limitations of the technology and the potential impact of AI on established practices in NPD. The article establishes a research agenda for exploring the use of language models in this area and the role of humans in hybrid innovation teams. (Note: Following the idea of this article, GPT-3 alone generated this abstract. Only minor formatting edits were performed by humans.)

Practitioner points

Transformer-based language models like GPT-3 are a powerful type of AI that can perform various tasks during an innovation process like text summarization, sentiment analysis, insight generation, or idea generation at an incredible scale.

Such technologies support the exploration of larger problem and solution spaces and can augment humans to improve innovation performance.

Artificial intelligence and humans will increasingly work together in a form of hybrid intelligence, which calls for a re-evaluation of how we approach and manage innovation.

1 INTRODUCTION: AI AND INNOVATION

Over the last few decades, we could observe numerous technological advancements in the field of artificial intelligence (AI), including algorithmic breakthroughs, benefits of digitalization in the form of inexpensive data collection and handling, open-sourcing of key technologies, and access to cloud-based services (Bleier et al., 2020; von Krogh, 2018). These advancements have increased the adoption of AI in a variety of industries and led to significant productivity advancements in operational business functions. They could also enable a more data-driven and AI-based approach to innovation (Cockburn et al., 2018; Kakatkar et al., 2020) that could bear new possibilities to dramatically increase the productivity of new product development (NPD)—an endeavor much more challenging than increasing the efficiency of a production task (Bloom et al., 2020; Brynjolfsson et al., 2019).

Among the different types of AI, transformer-based language models may bear particularly interesting opportunities for an AI-augmented innovation process. They are a type of AI designed to process and generate (natural) language in the form of text and can be used for various tasks such as machine translation and text summarization, insight extraction, or generating creative output. This makes language models an especially interesting form of AI in knowledge-intensive work such as NPD. Examples of such models include Google's BERT or OpenAI's Generative Pre-trained Transformer (GPT-3). They not only reached human-like language understanding abilities (Zhang et al., 2021) but also feature an intuitive way of user interaction based on natural language. Transformer-based language models have advanced so much that publishers now use them to write entire articles (The Guardian, 2021) or, for example, the Parliament of Finland (Eduskunta) (2021) includes them as a member of parliamentary panels to complement humans in debates. Furthermore, numerous companies have started to deploy such models to tackle innovation challenges (see, e.g., Toews (2022) for an overview of companies providing products and services based on transformer-based language models). Examples include "Hebbia", which built a search engine to tap into vast amounts of unstructured data and extract relevant insights; "Inceptive", which applies language models to the development of RNA therapeutics and vaccines, or "CopyAI", which uses the generative abilities of transformer-based language models to automate writing a marketing copy.[1] These examples already indicate the opportunities to use this class of AI for automation and efficiency improvements along the stages of the innovation processes.

At the same time, transformer-based language models ask us to reconsider many established practices in NPD. For example, their creative abilities question common practices of how good ideas are created, like using creativity techniques or other divergent thinking tools. Specifically, the role of humans in NPD will change as AI takes on more tasks. So how will innovation teams deal with AI-generated and AI-co-created ideas? How do we address the fear of human experts being replaced by a machine? Will designers be reduced to "prompt engineers", that is, just designing the task given to an AI instead of the product or component? Augmenting human innovation teams with AI hence calls for investigating many aspects of jointly working with AI rather than only human colleagues.

With this catalyst, we hope to spark a discussion on how transformer-based language models will impact innovation and how this affects extant research in the field of NPD. The following section will briefly introduce the basic technology behind transformer-based language models. We then present the AI-augmented double diamond as a framework to guide our exploration of transformer-based language models for innovation. Next, we discuss several concrete examples of utilizing these models in typical tasks of an innovation project, using the GPT-3 algorithm, a powerful and easily accessible transformer-based language model. We use these examples to develop a set of research questions that arise from integrating such AI into NPD teams, also discussing the limitations of such technologies. Our goal is to spark further debate and research on the opportunities, limitations, and managerial implications that emerge when artificial and human intelligence are combined to solve complex innovation tasks in a way that none of them alone could have accomplished.

2 TECHNICAL BACKGROUND: LANGUAGE MODELS

Transformer-based language models are a special kind of AI used for natural language processing (NLP), which is a range of computational techniques for analyzing and representing naturally occurring texts to achieve human-like language processing (Liddy, 2018). In general, NLP is not new to innovation management. Previous research has explored such techniques in the area of text analysis (text mining) and how they can be applied to innovation processes (Antons et al., 2020). Whereas earlier models have typically been very task-specific, newer NLP technologies can take on multiple innovation-related tasks. Especially so-called generative or transformer-based language models have recently moved to the center of attention. Many transformer-based language models are autoregressive models (e.g., GPT-3), meaning they predict a word based on its preceding words in a text. NLP has seen continued progress over the past decades with a trend toward larger and more complex models rapidly increasing their capabilities—from the mere suggestion of related words (Garay-Vitoria & González-Abascal, 1997) to state-of-the-art models that can produce full newspaper articles indistinguishable from human-written text (Brown et al., 2020). For readers interested in technical details and developments leading to transformer-based language models, Table 1 presents an overview of important technical milestones in this field.

TABLE 1. Selected scientific milestones in natural language processing (NLP)

Reference	Title	Milestone
Garay-Vitoria and González-Abascal (1997)	Intelligent word-prediction to enhance text input rate (a syntactic analysis-based word-prediction aid for people with severe motor and speech disability)	Introduces word-prediction based on a syntactic analysis
Bengio et al. (2003)	A neural probabilistic language model	Introduces neural networks for calculating probability functions for word sequences
Mikolov et al. (2010)	Recurrent neural network (RNN) based language model	Introduces RNNs for language modeling
Graves (2013)	Generating sequences with RNNs	Introduces long short-term memory (LSTM) RNNs for text prediction
Bahdanau et al. (2014)	Neural machine translation by jointly learning to align and translate	Introduces attention mechanisms
Vaswani et al. (2017)	Attention is all you need	Introduces the transformer architecture
Devlin et al. (2018)	Bert: Pre-training of deep bidirectional transformers for language understanding	Introduces large bi-directional transformer-based language models
Radford et al.	Improving language understanding by generative	Introduces a line of generative

The "Transformer Architecture" introduced by Vaswani et al. (2017) provides the basis for most state-of-the-art language models. These models can take the context of the processed words into consideration, which allows for a more nuanced understanding of related words and concepts. Because the capabilities of these models significantly improve with model size (Brown et al., 2020; Radford et al., 2019; Tamkin et al., 2021), the rapid increase in model sizes—newer versions, such as GPT-4, are expected to have 100 trillion parameters—has severe implications for the usefulness and applicability of transformer-based language models for innovation tasks.

3 AI-AUGMENTED KNOWLEDGE-BASED PRACTICES

Knowledge is a key resource central to a firm's innovation activities (Crossan et al., 1999). A large body of literature has conceptualized innovation management as a knowledge-based practice (e.g., Chung & Lee, 2020; Silva et al., 2018). Such practices are inseparably connected to language—be it talking to a colleague, listening to a lecture, or reading an article.

Understanding language grants us access to a plethora of knowledge. This makes language models an especially interesting form of AI for knowledge-intensive work such as NPD.

The AI-augmented double diamond framework

To structure our exploration of potential use cases of how transformer-based language models can augment human innovation teams in the future, we refer to the double diamond framework (Figure 1, upper picture). Popularized by the design thinking community since the mid-2000s (Design Council, 2022), it builds on the works of Guilford (1956), who is generally attributed with originally making the distinction between divergent and convergent thinking. Innovation processes consist of a sequence of steps where innovators first explore a wide range of problems and opportunities to then decide on adequate (technical) solutions to the given problems (Marion et al., 2023; Marion & Fixson, 2019). In both stages, innovation teams perform divergent and convergent tasks. Successfully navigating through these processes requires building on substantial amounts of existing knowledge and ultimately generates new knowledge. Building on the conceptualization of innovation processes as a double diamond, we present the AI-augmented double diamond (Figure 1, lower picture) as a framework that highlights how AI and especially transformer-based language models can be particularly useful for fostering divergent processes and help to explore larger problem and solution spaces.

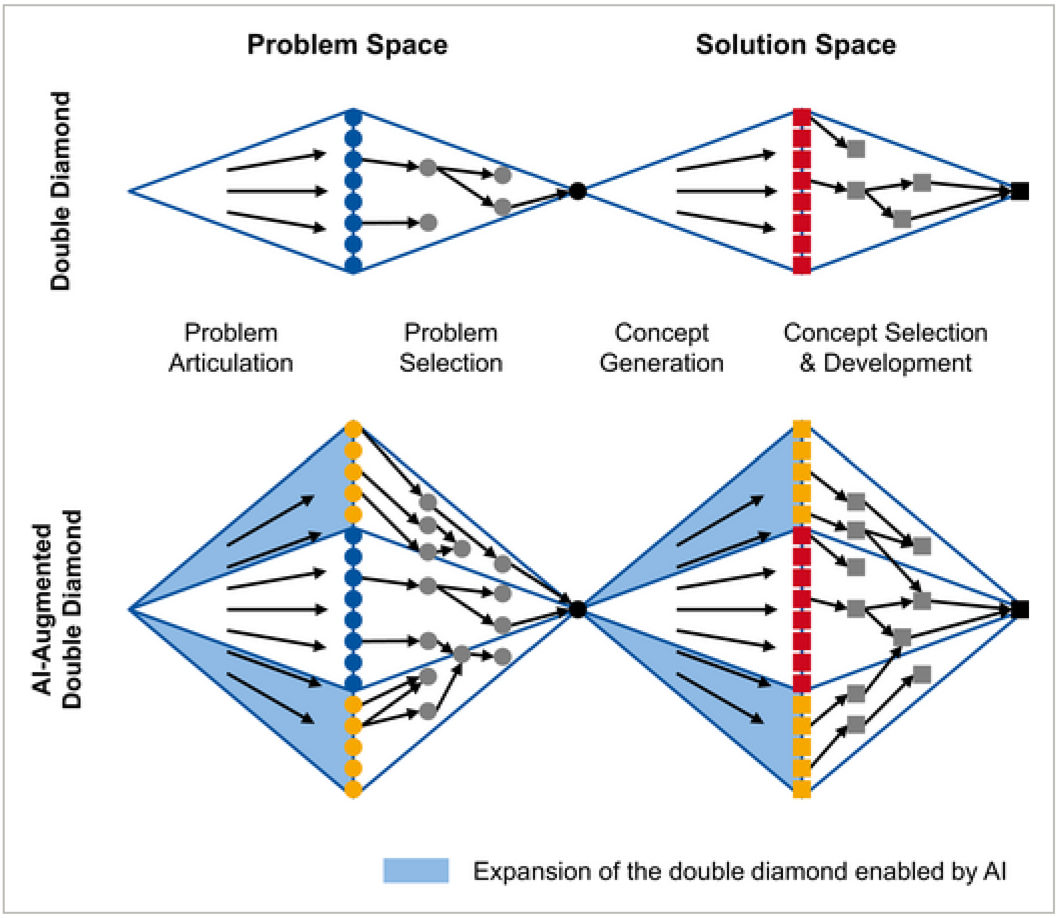

FIGURE 1

Open in figure viewer | PowerPoint

The original double diamond framework (above), as conceptualized by Marion and Fixson (2019), and the artificial intelligence (AI)-augmented double diamond framework (below)

By expanding the problem and solution spaces in which NPD teams can operate, language models create an opportunity to access and generate larger amounts of knowledge, which in turn results in more possible connections of problems and solutions. This should ultimately lead to qualitatively superior solutions and higher innovation performance. In the following, we discuss specific applications of transformer-based language models in NPD, highlighting their potential to improve innovation productivity by augmenting human innovation teams in exploring larger problem and solution spaces. For our exploration of exemplary use cases, we utilize the GPT-3 algorithm. This language model has fundamentally shaped the AI space in recent years (Zhang et al., 2021). It was first introduced in 2020 and is offered by OpenAI, a Silicon Valley-based startup founded in 2015. In December 2022, many media reports about ChatGPT, a chatbot built on top of GPT-3 and fine-tuned with supervised and reinforcement learning techniques to generate detailed responses on a large variety of topics, made GPT-3 known to a larger audience. For readers interested in replicating our use cases, we provide the technical parameters used to tune the model in a web appendix to this article (Appendix A1 has more information on how to access and interact with this AI model).

Exploring problem spaces with AI

Exploring problem spaces includes building on existing knowledge to identify opportunities for innovation. To do so, innovation teams employ knowledge-based practices such as knowledge capitalization and knowledge spanning, which, at their core, help to access knowledge that exists either within an organization or in its network (Silva et al., 2018). To make such existing knowledge usable, innovators have to extract knowledge that might be coded explicitly or implicitly in a given knowledge base. Members of an innovation team can then further process the extracted information, internalize it, and convert it into new knowledge that their firm can then use in their innovation processes.

Typically, extracting vast amounts of knowledge is rather labor-intensive and not easily scalable because much knowledge is encoded in unstructured data that is still difficult to process automatically (Fan et al., 2012) and, therefore, involves humans manually reading and extracting information and knowledge. Earlier machine learning (ML) models have already automatized parts of such processes. ML algorithms perform well on a variety of pattern recognition tasks that are relevant for knowledge extraction. These can range from detecting patterns in visual data, for example, for quality control or analyzing technical samples of an experiment, to identifying novel ideas in online communities (Christensen et al., 2017) or customers with lead user characteristics (Kaminski et al., 2017). Transformer-based language models like GPT-3 are the next level of technology for such knowledge extraction practices. Their flexibility and generative capabilities provide ample opportunity for different knowledge extraction practices, allowing NPD teams to apply one model for a large variety of tasks. Their context awareness plays a critical role in understanding important connections within a given text and extracting relevant information and knowledge. In the following, we provide three specific examples, namely text summarization, sentiment analysis, and customer insight generation, which highlight different knowledge extraction capabilities of one state-of-the-art transformer-based language model. All of which can help to diverge and explore larger problem and solution spaces in the double diamond framework.

3.2.1 Text summarization

Extracting knowledge from text requires the identification, selection, and processing of large amounts of textual data, which can be very costly and time-consuming as well as overwhelming for humans. This might lead to unwanted consequences such as biased decision-making, as humans tend to make simplifying assumptions to be able to process such large volumes of information (Simon, 1955). Transformer-based language models provide an adequate solution to such problems through their ability to summarize text efficiently while conveying the most important aspects and meaning of the original text. They are especially well suited for such tasks as to their context awareness and the ability to pay special attention to the most relevant pieces of information. This gives them a great advantage over earlier text-summarization algorithms, which struggled to understand sentiment and meaning in textual data (Fan et al., 2006). Nonetheless, we also acknowledge that algorithmic approaches to text summarization are not without potential trade-offs, such as questions regarding the trustworthiness of AI-generated summaries. The potentials and limitations of this technology raise important questions for further research as well as practitioners that we address later in this article.

To showcase the text summarization capabilities of state-of-the-art transformer-based language models, we provided GPT-3 with the text of a blog post reviewing camping gear (in this particular case a portable power station). This serves as an example of a source of potentially valuable information for innovation projects. At the time of writing this article, a Google search for "camping blog" produced 299 million search results, far more than any human could ever read when building background knowledge for, for example, a product development project in this domain. Therefore, having a way of extracting the most relevant information from such articles through AI-generated text summaries can provide an easy solution to identifying the most relevant articles and exploring a larger problem space. Having chosen an exemplary blog post, we first prompted GPT-3 to summarize the text (Prompt A2.1 in Table A2-1 in the web appendix) and then prompted the model to summarize the text in just one sentence (Prompt A2.2) to test an extreme case of automated summarization.[2] This is the summary created of a long blog post by the GPT-3 algorithm (as a response to Prompt A2.1):

"

> The Jackery Explorer 500 is a portable power station that can be charged up in a variety of ways and then used for your off-grid adventures. It has a capacity of 518 Wh (21.6 V, 24 Ah), a cell of Li-ion NMC, and a lifecycle of 500 cycles to 80%+ capacity. It can be recharged in 7.5 h via AC adapter, 8 h via 12 V car adaptor, or 9.5 h via Jackery Solar panel (sold separately). It has multiple outputs, including AC (230 V, 500 W, 1000 W surge), USB-A (5 V, 2.4A), car (12 V, 10A), and DC (12 V, 7A). It is designed to be used for charging devices such as phones, cameras, laptops and drones, and can also be used for mini-coolers and TVs. It is priced at £529.99 for the power station alone, or £749.99 with the solar panel included.

This and the other examples documented in the web appendix show the general ability of transformer-based language models to capture the essence of an original text and reproduce a coherent, yet substantially summarized version of it. In this way, language models help to speed up knowledge extraction by supporting humans in quickly understanding the essence of a text and identifying relevant text passages faster. Hence, innovation teams can reduce knowledge-extraction efforts and re-distribute their focus to applying the acquired knowledge to their innovation tasks. In addition, transformer-based language models can help overcome problems of myopic search by enabling users to access more data more quickly that might otherwise be outside their area of expertise.

Despite these potential upsides of using such technologies, it should be noted that there currently is no way of telling how reliable an AI-generated response is without knowing the answer to a question oneself. This is somewhat defeating the purpose of using an AI in the first place. Many models are currently able to generate very plausible text passages that sound as though they could be true, but might not necessarily be. Therefore, a weakness of transformer-based language models could be that they do not just retrieve answers to a question from a set of pre-defined responses, but generate an original response from scratch. If provided with the task of generating output, a language model will not tell a user that it does not know the answer (as a human hopefully would do)—it will simply produce the most likely answer. In this sense, transformer-based language models are still a rather narrow form of AI, even if they bring us one-step closer to reaching artificial general intelligence.

Problems of reliability arise specifically in areas where rather limited amounts of relevant knowledge on a subject were included in the initial training data. Innovation generally happens at the frontier of knowledge and oftentimes requires firm-specific knowledge that is not encoded in text corpora used for training of general language models. However, companies can potentially mitigate this problem by re-training or fine-tuning existing models with their firm-specific knowledge. Nonetheless, humans using such technology to innovate need to be aware that such limitations exist and have to use their expertise to screen AI-generated output for such problems. Hence, further research needs to investigate in which contexts and under which conditions transformer-based language models can perform a text summarization in a better way than humans. But rather than a comparison of humans vs. machines, further research should investigate how humans' domain expertise can support language models to utilize their abilities best, for example, by formulating the summarization task in a specific way or pre-selecting the text corpus for analysis.

3.2.2 Sentiment analysis and insight generation

Another task that transformer-based language models are well equipped to handle is to mine the overall sentiment of a given text corpus. Especially online communities contain information that is valuable for innovating firms, as customers discuss products, services, and trends online that have an impact on a firm's product development process (e.g., Blazevic & Lievens, 2008). For instance, customer reviews contain important information on whether customers liked or disliked certain product features or what needs might not have been met. Sentiment analysis is an established methodological approach for analyzing data from social media streams, user forums, or customer reviews (Feldman, 2013). It is valuable as it helps firms to better understand customer needs and can highlight areas where improvements to products and services are needed. As extracting relevant sentiment from large streams of data manually is a laborious task, sentiment analysis has been automatized through ML approaches before. However, similar to text summarization, this required special software and often costly licenses. Broad transformer-based language models like GPT-3 hence promise an opportunity to greatly improve the accessibility and scalability of sentiment analysis.

To explore these abilities, we took a set of 10 customer reviews for an electric portable air pump from Amazon.com.[3] We chose five reviews with positive and five with negative sentiment, specifically selecting reviews that varied in length and writing style to provide a realistic real-world sample. Provided with just four examples containing our own (human) assessment of whether the sentiment was overall positive or negative, GPT-3 labeled all of the remaining six customer reviews correctly, showing that it can understand sentiment within text.[4] Again, the model was originally not specifically trained to perform sentiment analysis. While we only provide 10 customer reviews in this example, this approach is highly scalable and can easily be applied to basically infinitely many reviews (or other forms of text) showing how this can help to cover a larger problem space than would be feasible for humans alone.

A typical insight for an innovation project is not just information on whether customers respond positively or negatively to a given product, but rather understanding what customers specifically liked or disliked. Hence, we built on the previous example to test GPT-3's question-answering capabilities. Switching tasks mid-prompt, we continued the session outlined before by asking GPT-3 to identify features of the product customers (dis)liked (Prompt A2.7 in Table A2-3 in the web appendix). The model understood that the task had changed and was able to correctly answer which features customers (dis)liked. It also specified in its response why customers (dis)liked these features. This is the output of the model (we present another example and model parameters in Table A2-3 of the web appendix):

> "The most common feature that customers liked was the built-in light. The light was bright enough to illuminate a tent or a picnic table, and it was a convenient way to light up a dark campsite."

> "Sentiment: Positive."

This example highlights the flexibility of transformer-based language models. In a way, these models possess the ability to understand customer needs and can extract such knowledge with great accuracy and without extensive training—and at a scale that humans are not able to cope with efficiently (just consider the millions of reviews added to a site like Amazon every single month). Automatically distilling such knowledge to a level where humans can then continue to work with helps innovation teams to concentrate on turning the insights collected during the problem articulation phase into value-creating new products. These opportunities for augmenting human innovation teams raise questions about whether conventional approaches to market intelligence or social media analysis in NPD should still be used, and when.

Further research needs to test and validate these opportunities and potential constraints, also considering how humans and AI interact with and inform each other. Consider our example of providing our human perceptions of "positive" and "negative" in the few-shot learning process above (also see Prompt A2.6 in the web appendix). How dependent is the algorithm's outcome on a human's perception of positive or negative? Does an algorithm perceive, for example, a negative comment still as constructive, providing valuable insights, while a human might have dismissed such a comment as "just another complaint"? This raises many questions about how tasks should be allocated to humans or AI in the innovation process of the future, which antecedents influence this division of labor, and under which circumstances a collaboration of humans and AI provides the best results. We will explore the latter thought of a "hybrid intelligence" (Dellermann et al., 2019) in the discussion section of our article.

Exploring solution spaces with AI (idea generation)

After exploring problem spaces and extracting relevant pre-knowledge, innovators explore solution spaces and thereby generate new knowledge previously not available to their organization. Such knowledge-based practices are at the core of our understanding of problem-solving and innovation (Crossan et al., 1999). Corresponding processes generally involve innovation teams employing their creativity; a trait that for a long time had been considered inherently human (Boden, 1998). However, there is growing potential in the application of AI to creative tasks (Amabile, 2020). Especially transformer-based language models can be highly relevant in this context due to their ability to generate ideas based on a question or other initial input, as some practitioners such as the design firm IDEO have started to explore (Syverson, 2020). Because of their few-shot learning capabilities, they can generate adequate responses to a given problem statement and come up with original and useful ideas when prompted with just a few examples of what typical brainstorming results look like. In addition, users of such models can precisely tune them to produce more creative (radical) or more deterministic (incremental) responses—an ability rarely imaginable for humans. While we focus on idea generation in this article, transformer-based language models can also generate other text-based output, such as computer code, which plays an important part in many NPD projects today. [5]

An important activity at the front-end of innovation, but also during technical development, is idea generation. It has been shown as being central to a firm's innovative performance (Kijkuit & Ende, 2007). To generate valuable ideas, we employ creativity, generally defined as "the production of novel, useful ideas, or problem solutions" to a given task (Amabile et al., 2005: 368). To explore transformer-based models in idea generation, we first continued the camping gear example from the previous section. We asked GPT-3 to "Create a list of ideas for ways to improve an air pump made for camping.", and it came up with ideas like adding a carrying handle to the air pump for easy transport or adding an adapter for inflating inflatable camping mattresses (see Table A3-4). Building on one of these ideas, we prompted GPT-3 with a new formulation of the task ("How might we make a portable air pump more useful for campers?"), and now also provided one example idea (to micro-train the model). The algorithm came up with a new set of ideas, which all seem to be reasonable and possible starting points for new features in an NPD project (see Table A3-5 in the web appendix). While humans could also have generated these ideas easily, the algorithm produced them within seconds, providing a continuous flow of ideas whenever prompted, but fine-tuning an idea further when prompted to do so. This can largely increase the creative solution space of an NPD team, at almost no cost.

We also explored a more open (fuzzy) question, tasking GPT-3 with exploring opportunities of a new technology. We prompted the AI with two similar, but slightly different questions regarding the use of transformer-based language models for innovation, providing in both cases four ideas generated by us as a starting point of the human-AI interaction (Table A3-1 in the web appendix). The AI-generated ideas contain a few notable characteristics that highlight how language models can prove valuable in an innovation project. On the task "How might we use transformer-based language models in innovation processes," the algorithm suggested to ...

"

Use language models to generate ideas for patent applications.

Use language models to find similar patents to help in patent searching.

Use language models to help in monitoring communication between different departments or people.

Use language models to predict trends in what people search for in the future.

Use language models to generate possible headlines for news stories.

Use language models to generate possible names for new products.

"

Except for the second-last response, which only loosely relates to innovation processes, all responses generated by the AI are relevant in the given context and are viable suggestions. While some answers are more original than others, they build on the four provided examples without just copying them. As the responses are both useful and novel, we can induce that transformer-based language models such as GPT-3 are capable of producing creative output. This is exactly what Amabile (2020: 251) calls AI creativity: "the production of highly novel, yet appropriate, ideas, problem solutions, or other outputs by autonomous machines."

Another interesting observation in our exploration stems from the comparison of the AI-generated ideas for prompts A3.1 and A3.2 (in Table A3-1 in the web appendix). For both tasks, the examples provided by us to train GPT-3 were the same. We only slightly modified the task (with no changes to the model's settings), asking for opportunities "to foster innovation" in the second case. Still, the differences in the results are striking. For the slightly more general prompt A3.1, the model generated ideas related to different stages in an innovation process – from the front-end to the diffusion stage. The response to the more directed prompt A3.2 covered various stakeholders involved in innovation processes, from taking a micro-level focus on team members, an organization-level focus on how firms can use the technology to improve or expand products, to a customer-focused perspective with use cases to make products more useful and fun. Hence, the model seems to be able to take different perspectives. This suggests that GPT-3 implicitly learned during its initial meta-learning process that there are different viewpoints on managing innovation. Remember that we only prompted the model with the word innovation, but did not provide any more explanation or context on what innovation is, how it works, and what is relevant in this context.

These examples show the impact that "prompt engineering" can have on the output of such AI models. Prompt engineering recently evolved as a way humans work with complex AI systems, particularly NLP tools like GPT-3. The more autonomous machines become in providing the desired outcome, the more critical the task that is promoted to the machine producing the outcome. While prompt engineering has been covered in the AI and engineering literature, we still have no knowledge about this new function in innovation management – but perhaps can build on the rich kinds of literature on task and problem formulation in creativity workshops or open innovation contests.

Related to this, we explored how changing model parameters like the temperature setting (see Table A3-2 in the web appendix) can be used to control the scope (openness) of the search space (performing a narrower or broader search). Potentially, this could be used to adapt to changing goals within an ideation session. Given the large body of research on openness of search in innovation management (Salter et al., 2015; Sofka & Grimpe, 2010), this ability could have a large influence on how firms approach open innovation in the future, as it makes it far easier to tap into external knowledge that is encoded in the AI's knowledge base. Generally, the sheer amount of data used for training purposes of transformer-based language models can provide the basis for a very broad exploration of relevant ideas from various contexts and should be able to support divergent thinking.

However, the examples also highlight that there is still room for improvement. Some of the ideas are not very elaborated. Still, humans in an innovation team can use the ideas generated by the language model and build on them. They can feedback these results into the model as new examples during further few-shot learning, generating new AI-generated responses by iteratively building upon new knowledge—exactly as humans would do in a brainstorming workshop. But how exactly such hybrid iterations between humans and AI in the innovation process can be facilitated best remains the subject of further research. Especially looking into the organizational antecedents of successful AI-augmented innovation projects could be a promising research avenue. Research could investigate how the implementation and adoption of transformer-based language models are fostered by dedicated structures and processes, including governance decisions like using open source or proprietary language models or how to place the AI capabilities in an innovation team. This also asks for a reconsideration of established frameworks of knowledge absorption and adoption, for example, considering the Not-Invented-Here syndrome (e.g., Antons & Piller, 2015) or different judgments of errors and failures made by humans and algorithms.

An important limitation to consider here is that the original training data for language models has a cut-off point after which new knowledge is no longer contained in the training set used for unsupervised learning. Hence, critical information might not be included in the model's knowledge base. Users have to be conscious of this aspect when interacting with a model. Generally, this natural data cut-off point calls for a continued re-training of models in use. While this aspect might not be critical for applications such as lyric composition or the writing of novels, it is especially relevant in the innovation sphere as innovators should base their decisions on the newest knowledge available. This aspect is even more pronounced in research fields where this knowledge stock expands rapidly. At the same time, the re-training of these models is very easy. These models can very effortlessly acquire new knowledge that can then be incorporated into innovation processes, as long as the information is available in a machine-readable form.

Another important aspect is the quality of the data used for training purposes. Just as with any other type of AI that is based on large amounts of data, language models can potentially suffer from biases. Most models use training data based on text crawled from the internet. Such text generally has not specifically been checked for bias. Developers of models like GPT-3 seem to be aware of this problem and discuss gender, race, and religious biases in their article (Brown et al., 2020). In addition, there are active attempts to increase the quality of the databases used for training. For example, Radford et al. (2019) have created the WebText dataset that only includes text from websites whose outbound links on Reddit have received at least three karma on the platform.[6] Such measures can ensure a certain degree of data quality.

However, the bias problem is still far from being solved (Heaven, 2020). Therefore, users of language models should always be aware that they could generate biased content. Consequently, employing language models in teams with humans that can check for potential bias appears to be a sensible approach to incorporating such AI in innovation processes—underlying the idea of hybrid intelligence. This also necessitates the active management of such teams, including, for example, the training of employees in searching for and reflecting on biases. Acquiring such skills might then even help humans become more aware of their own biases, allowing humans and AI to learn from each other.

4 DISCUSSION: THE EMERGENCE OF HYBRID INNOVATION TEAMS

Exploring several use cases, we showed how a specific type of AI, transformer-based language models, can help NPD teams to extract meaningful insights and sentiment from large volumes of text and can contribute to ideation and problem-solving tasks along the entire double diamond model. The ability of transformer-based language models to integrate knowledge from various and diverse knowledge bases allows innovators to increase the quantity, quality, and diversity of ideas and to create more value in their innovation processes—at a very low cost (the price we had to pay OpenAI for executing all tasks explored in this article was less than $1). Consequently, AI blurs the line between internal and external knowledge and allows innovation teams to tap into a larger pool of knowledge than was previously possible.

In this respect, language models could be seen as a further iteration of the open innovation paradigm. The abilities of language models to interact with different knowledge sources, learn from them, and share and transform knowledge allow this kind of AI to act as a knowledge broker that facilitates sharing of knowledge between different stakeholders, while also fostering the creation of new knowledge (Waardenburg et al., 2021). Thereby, AI can take on some of the functions that were traditionally provided by innovation intermediaries. Through their flexibility and adaptability, enabled by meta-learning during pre-training, innovation teams can employ these language models to access existing knowledge outside a firm. Models that have been trained on large text corpora from the internet have knowledge on a wide range of topics, which opens up the opportunity for innovation teams to integrate knowledge that might lay outside their area of expertise. Given a prompt by a human, AI can help to establish connections between concepts and ideas that might otherwise not have been obvious to humans alone. Few-shot learning capabilities then allow for easier interaction between an innovation team and the AI. The team members only have to provide a limited amount of exemplary responses to a given task, so that the language model can generate adequate output. A human actor can then build upon this output. By providing the language model with their own knowledge and ideas through natural language, innovation teams can integrate the AI into their existing processes, as if it would be a new colleague. The combination of knowledge from human team members and knowledge provided by AI provides the opportunity to greatly improve the productivity of NPD practices and to produce outcomes that would not have been possible with just the skillset of one of the actors.

Our exploration not only demonstrated the capabilities of these algorithms, but also identified many opportunities for future research focusing on particular abilities and applications of AI models for the various stages of the innovation process, like opportunity recognition (e.g., identifying trends and customer needs, predicting technology trends, or providing technology forecasts), ideation and concept development, concept selection (e.g., selection algorithms overcoming human decision biases), design and development (e.g., generative design algorithms to create technical designs, prototypes, and solutions), and launch and lifecycle management (e.g., approaches to mine large pools of information on product performance or customer satisfaction). Considering the fast-paced development of AI, we expect even more powerful models soon.

Still, it is unlikely that these models will be working autonomously, that is, that new products can be developed entirely by a set of algorithms interacting with each other, with humans just providing the initial prompt. Our exploration of the GPT-3 algorithm for various use cases suggests that we consider algorithms rather as members of an innovation team. Hence, we propose that transformer-based language models will specifically support knowledge-based innovation practices in the form of hybrid intelligence, that is, the combination of human and artificial intelligence, "thereby reaching superior results to those each of them could have accomplished separately, and continuously improve by learning from each other" (Dellermann et al., 2019: 640). Hybrid intelligence aims to combine the best of both worlds (Piller et al., 2022). Some tasks in the innovation process can be probably automated by an algorithm more reliable and cost-effective than they can be conducted by humans. An example is the continuous scanning of social media posts to derive latent customer needs (opportunities) in the front-end of NPD. However, humans still have unique capabilities. Consider situations that need empathy, creativity, or ethics (van der Aalst, 2021). Decisions here demand human contributions and cannot entirely be executed by a machine. AI and human intelligence will complement each other.

Hence, building and orchestrating such hybrid teams and allocating tasks between humans and machines becomes a new management task for project leaders. Understanding the contingencies and success factors of these allocation decisions is a domain with a wide demand for further research. Managers will have to take into consideration the distinct characteristics of human and non-human actors and their collaborations. While humans will play a major role in providing context, steering language models toward desired results, and embedding AI output in the larger innovation picture, transformer-based language models can speed up many tasks that require the handling of large amounts of text, understand patterns in data invisible to humans, and make connections between knowledge bases that might not be readily available to human team members (Piller et al., 2022). At the same time, including algorithms as team members may also change the way how humans interact with each other in a team. In the concluding section of our article, we will discuss some of the related research questions in more detail.

5 CONCLUSION AND OPEN RESEARCH QUESTIONS

Transformer-based language models are a powerful form of AI that can augment human innovation teams and hold the potential to fundamentally change how knowledge-based practices will be approached in the future, thereby improving innovation performance. Their ability of understanding, generating, and adapting language, fostered by meta- and few-shot learning and attention mechanisms, make these AI models a powerful tool in any innovator's toolkit, enabling innovation teams to explore larger problem and solutions spaces. However, transformer-based language models are not without limitations. We discussed their most important constraints throughout this Catalyst but want to emphasize that these limitations do not render transformer-based language models useless, but rather highlight that such technology should not be trusted and used blindly. Given the speed of development in this field, we are optimistic that future generations of language models will mitigate at least some of the limitations of current state-of-the-art models.

Comparing their skills and limitations, we consider transformer-based language models not as a stand-alone technology, but rather as an actor in an innovation team that needs to be integrated into existing processes in the understanding of hybrid intelligence. At present, we are far from understanding all the potential use cases, benefits, and pitfalls associated with such hybrid innovation teams. More research is needed, not only from a technical but particularly also from a managerial perspective. We encourage scholars to take our work as inspiration to dive deeper into the possibilities and implications for innovation management associated with transformer-based language models. Hence, we conclude this Catalyst with two sets of research questions. Table 2 contains those questions we find especially relevant, considering our own practical experience of working with GPT-3 and reflecting on the results. This list is by no means exhaustive but can provide a starting point to explore the various aspects that transformer-based language models may affect, including units of analysis on the technical, individual, team, organizational, and industry (meta) level. In addition, building on the idea of this article, we also prompted the GPT-3 algorithm to continue our list of research questions presented in Table 2. Table 3 provides its output. We hope that both sets of questions inspire other scholars to investigate transformer-based language models and their implications for NPD and spark a fruitful discussion on the collaboration of humans and AI for innovation.

TABLE 2. Research themes on transformer-based language models for innovation

Research theme	Level of analysis	Research question
AI capabilities	Technical	What are the differences between problems and solutions identified by an artificial intelligence (AI) and by humans from a given body of text? How and when are humans still superior in reading "between the lines" and making serendipitous discoveries?
		How do different prompts and task formulations influence the outcome of transformer-based language models? How can we formulate problem statements that are best suited for the usage in AI models for different stages of new product development (NPD)?
		How and when is AI better equipped to pick up on relevant ("sticky") information that humans might overlook? Can an AI identify latent need information?
		(How) can an AI prioritize and select high-quality ideas?
Role of humans/AI in hybrid innovation teams	Team and Individual	Which factors influence how work shall be allocated to humans versus AI models in hybrid innovation teams? What is the right balance between automation and human collaboration?
		How will the introduction of AI to processes like idea generation change the role of humans in creative tasks?
		(How) will humans and AI jointly define search spaces? How will humans and AI interact and inform each other?

TABLE 3. Research questions generated by GPT-3 (classification added by authors)

Research theme	Level of analysis	Research questions
Managing innovation	Organizational	What are the best ways to manage and monitor AI-based innovation processes?
	Process	How does the increasing use of AI in innovation processes change the way we think about innovation?
Effects on society and the economy	Meta	What are the best ways to measure the impact of AI on innovation processes?
		What policies or ethical considerations need to be in place to ensure that AI-based innovation processes are responsible and beneficial to society?
		What are the possible long-term effects of AI-based innovation processes on society and the economy?
		What are the potential risks and benefits of AI-based innovation?
		How can we ensure that AI-based innovation is responsible and beneficial to society?

Abbreviations: AI, artificial intelligence; GPT-3, Generative Pre-trained Transformer 3.

Answering such research questions requires scholars to apply a multitude of research methods and interdisciplinary research. We still lack dedicated theories that can explain and design collaborative ideation and problem-solving by humans and AI. Hence, answering these questions will often best be achieved through inductive and abductive, rather than deductive reasoning (Amabile, 2020; Bamberger, 2018; von Krogh, 2018). Following a grounded theory approach and building on a multitude of research methods is likely to yield the best results. First, observing ensembles of humans and AI in a real-world context could lead to valuable insights into possible tensions arising in such collaborative environments. One example that lends itself well to such an approach is the use of generative design software by NPD teams. A similar approach has been successfully applied to studying the introduction of a predecessor of generative design software, namely early CAD software, into NPD processes (Thomke, 1998). Accordingly, such an approach would mostly rely on observations, interviews, and supplemental material provided by industry experts. Then, building on initial insights from such studies warrants the use of dedicated experiments to afford researchers more control over specifics of human-AI collaboration in innovation and allows for more targeted hypothesis-testing derived from observations made in a real-world context. Such experiments could, for example, test the performance effects of different constellations of hybrid intelligence, but also explore how prompt engineering would influence the outcomes or the acceptance of contributions by an AI to an NPD project.

ACKNOWLEDGMENTS

The authors thank the JPIM editors-in-chief for their guidance and advice on earlier versions of this article. The research has been funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy (EXC-2023 Internet of Production, funding ID 390621612) and the German Federal Ministry of Economic Affairs and Climate Action (BMWK) as part of the project SPAICER (funding ID 01MK20015A). Open Access funding enabled and organized by Projekt DEAL.

ETHICS STATEMENT

The authors have read and agreed to the Committee on Publication Ethics (COPE) international standards for authors.

Biographies

Sebastian Bouschery, M.Sc. is a doctoral student at the Institute for Technology & Innovation Management at the RWTH Aachen University. He received a Bachelor's and Master's degree in Business Administration and Mechanical Engineering from the RWTH Aachen University. His research focuses on the collaboration of humans and artificial intelligence in a form of hybrid intelligence. Particularly, he investigates how artificial intelligence can augment human innovation teams in creative innovation processes and how this affects innovation management.

Vera Blažević is an associate professor of Marketing at the Radboud University Nijmegen and visiting professor at the RWTH Aachen University at the Institute for Technology & Innovation Management. Her research interests include stakeholder co-creation, social processes in digital innovation management, and responsible innovation. Her prior work has been published in *Journal of Product Innovation Management*, *Journal of Marketing*, *Journal of Service Research*, *Journal of the Academy of Marketing Science*, *Industrial Marketing Management*, and *Journal of Interactive Marketing*, among others. She obtained her Ph.D. from the Maastricht University.

Frank T. Piller is a professor of management at the RWTH Aachen University, where he co-directs the Institute for Technology & Innovation Management. His research interests include open and user innovation, mass customization, managing disruptive innovation, and implications of new technologies for new product development. Frank's research has been published in *Journal of Product Innovation Management*, *R&D Management*, *Research Policy*, *Academy of Management Perspectives*, *Journal of Operations Management*, and *MIT Sloan Management Review*, among others. Frank obtained a Ph.D. in Operations Management from the University of Würzburg and previously worked at TU Munich, HKUST, and MIT.

Supporting Information

Filename	Description
jpim12656-sup-0001-supinfo.pdf PDF document, 410.3 KB	**Data S1.** Supporting Information.

REFERENCES

Amabile, Teresa M. 2020. "Creativity, Artificial Intelligence, and a World of Surprises." *Academy of Management Discoveries* **6**(3): 351– 4.

Amabile, Teresa M., Sigal G. Barsade, Jennifer S. Mueller, and Barry M. Staw. 2005. "Affect and Creativity at Work." *Administrative Science Quarterly* **50**(3): 367– 403.

Antons, David, and Frank T. Piller. 2015. "Opening the Black Box of "Not Invented Here": Attitudes, Decision Biases, and Behavioral Consequences." *Academy of Management Perspectives* **29**(2): 193– 217.

Antons, David, Eduard Grünwald, Patrick Cichy, and Torsten Oliver Salge. 2020. "The Application of Text Mining Methods in Innovation Research: Current State, Evolution Patterns, and Development Priorities." *R&D Management* **50**(3): 329– 51.

Bahdanau, Dzmitry, Kyunghyun Cho, and Yoshua Bengio. 2014. "Neural Machine Translation by Jointly Learning to Align and Translate." *ArXiv* **1409**: 0473.

Bamberger, Peter A. 2018. "AMD—Clarifying What we Are about and where we Are Going." *Academy of Management Discoveries* **4**(1): 1– 10.

Bengio, Yoshua, Réjean Ducharme, Pascal Vincent, and Christian Janvin. 2003. “A Neural Probabilistic Language Model.” *The Journal of Machine Learning Research* **3**: 1137– 55.

Blazevic, Vera, and Annouk Lievens. 2008. “Managing Innovation through Customer Coproduced Knowledge in Electronic Services: An Exploratory Study.” *Journal of the Academy of Marketing Science* **36**(1): 138– 51.

Bleier, Alexander, Avi Goldfarb, and Catherine Tucker. 2020. “Consumer Privacy and the Future of Data-Based Innovation and Marketing.” *International Journal of Research in Marketing* **37**(3): 466– 80.

Bloom, Nicholas, Charles I. Jones, John Van Reenen, and Michael Webb. 2020. “Are Ideas Getting harder to Find?” *American Economic Review* **110**(4): 1104– 44.

Boden, Margaret A. 1998. “Creativity and Artificial Intelligence.” *Artificial Intelligence* **103**(1–2): 347– 56.

Brown, Tom B., Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, and Amanda Askell. 2020. “Language Models Are Few-Shot Learners.” *ArXiv* **2005**: 14165.

Brynjolfsson, Erik, Daniel Rock, and Chad Syverson. 2019. “ Artificial Intelligence and the Modern Productivity Paradox: A Clash of Expectations and Statistics.” In *The Economics of Artificial Intelligence*, edited by Ajay Agrawal, Joshua Gans, and Avi Goldfarb, 23– 60. Chicago: University of Chicago Press.

Christensen Kasper, Sladjana Nørskov, Frederiksen Lars, and Scholderer Joachim. 2017. “In Search of New Product Ideas: Identifying Ideas in Online Communities by Machine Learning and Text Mining” . *Creativity and Innovation Management* **26**(1): 17– 30.

Chung, Euiyoung, and Kibaek Lee. 2020. “The Necessity of Anterior Knowledge Exchange Activities for Technological Collaboration and Innovation Performance Improvement.” *International Journal* *of Technology Management* **82**(1): 66– 96.

Cockburn, Iain M., Rebecca Henderson, and Scott Stern. 2018. " The Impact of Artificial Intelligence on Innovation: An Exploratory Analysis." In *The Economics of Artificial Intelligence*, edited by Ajay Agrawal, Joshua Gans, and Avi Goldfarb, 115– 46. Chicago: University of Chicago Press.

Crossan, Mary M., Henry W. Lane, and Roderick E. White. 1999. "An Organizational Learning Framework: From Intuition to Institution." *Academy of Management Review* **24**(3): 522– 37.

Dellermann, Dominik, Philipp Ebel, Matthias Söllner, and Jan Marco Leimeister. 2019. "Hybrid Intelligence." *Business & Information Systems Engineering* **61**(5): 637– 43.

Design Council. 2022. "What is the framework for innovation? Design Council's evolved Double Diamond." Accessed April 19, 2022. https://www.designcouncil.org.uk/our-work/skills-learning/tools-frameworks/framework-for-innovation-design-councils-evolved-double-diamond/
.

Devlin, Jacob, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. 2018. "BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding." *ArXiv* **1810**: 04805.

Fan, James, Aditya Kalyanpur, David C. Gondek, and David A. Ferrucci. 2012. "Automatic Knowledge Extraction from Documents." *IBM Journal of Research and Development* **56**: 1– 10.

Fan, Weiguo, Linda Wallace, Stephanie Rich, and Zhongju Zhang. 2006. "Tapping the Power of Text Mining." *Communications of the ACM* **49**(9): 76– 82.

Feldman, Ronen. 2013. "Techniques and Applications for Sentiment Analysis." *Communications of the ACM* **56**(4): 82– 9.

Garay-Vitoria, Nestor, and Julio González-Abascal. 1997. "Intelligent Word-Prediction to Enhance Text Input Rate." Proceedings of the 2nd International Conference on Intelligent User Interfaces, Orlando, FL.

Graves, Alex. 2013. "Generating Sequences with Recurrent Neural Networks." *ArXiv* **1308**: 0850.

Guilford, Joy Paul. 1956. "The Structure of Intellect." *Psychological Bulletin* **53**(4): 267– 93.

Heaven, Will Douglas. 2020. "How to Make a Chatbot that isn't Racist or Sexist." MIT Technology Review. Accessed December 24, 2021. https://www.technologyreview.com/2020/10/23/1011116/chatbot-gpt3-openai-facebook-google-safety-fix-racist-sexist-language-ai
.

Kakatkar, Chinmay, Volker Bilgram, and Johann Füller. 2020. "Innovation Analytics: Leveraging Artificial Intelligence in the Innovation Process." *Business Horizons* **63**(2): 171– 81.

Kaminski, Jermain, Jiang Yibin, Piller Frank, and Hopp Christian. 2017. "Do User Entrepreneurs Speak Different?" *Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems*. https://doi-org.ezproxy.auckland.ac.nz/10.1145/3027063.3053223

Kijkuit, Bob, and Jan Van Den Ende. 2007. "The Organizational Life of an Idea: Integrating Social Network, Creativity and Decision-Making Perspectives." *Journal of Management Studies* **44**(6): 863– 82.

Liddy, Elizabeth D. 2018. " Natural Language Processing for Information Retrieval." In *Encyclopedia of Library and Information Sciences*, edited by John D. McDonald and Michael Levine-Clark, 3346– 55. Boca Raton, FL: CRC Press.

Marion, Tucker J., and Sebastian Fixson. 2019. *The Innovation Navigator: Transforming your Organization in the era of Digital Design and Collaborative Culture*. Toronto, ON: University of Toronto

Press.

Marion, Tucker J., Mohsen Moghaddam, Paolo Ciuccarelli, and Lu Wang. 2023. " AI for User-Centered New Product Development: From Large-Scale Need Elicitation to Generative Design." In *The PDMA Handbook on Innovation and New Product Development*, 4th ed., edited by Ludwig Bstieler and Charles Noble. New York, NY: Wiley.

Mikolov, Tomas, Martin Karafiát, Lukas Burget, Jan Cernocký, and Sanjeev Khudanpur. 2010. "Recurrent neural network based language model." Proceedings of the 11th Annual Conference of the International Speech Communication Association, Makuhari, Chiba, Japan.

Parliament of Finland (Eduskunta). 2021. "Committee for the future heard AI probably as the first parliamentary committee in the world." Accessed December 26, 2021. https://www.eduskunta.fi/EN/tiedotteet/Pages/Committee-for-the-Future-heard-AI-probably-as-the-first-parliamentary-committee-in-the-world.aspx
.

Piller, F., Verena Nitsch, and Wil van der Aalst. 2022. " Hybrid Intelligence in Next Generation Manufacturing." In *Forecasting Next Generation Manufacturing*, edited by Frank Piller, Verena Nitsch, et al., 139– 58. Cham: Springer Nature.

Radford, Alec, Karthik Narasimhan, Tim Salimans, and Ilya Sutskever. 2018. "Improving Language Understanding by Generative Pre-Training." OpenAI Working Paper, San Francisco.

Radford, Alec, Wu Jeffrey, Rewon Child, David Luan, Dario Amodei, and Ilya Sutskever. 2019. "Language Models Are Unsupervised Multitask Learners." OpenAI Working Paper, San Francisco.

Salter, Ammon, Anne L. J. Ter Wal, Paola Criscuolo, and Oliver Alexy. 2015. "Open for Ideation: Individual-Level Openness and Idea Generation in R&D." *Journal of Product Innovation Management* **32**(4): 488– 504.

Silva, De, Jeremy Howells Muthu, and Martin Meyer. 2018. “Innovation Intermediaries and Collaboration: Knowledge–Based Practices and Internal Value Creation.” *Research Policy* **47**(1): 70–87.

Simon, Herbert A. 1955. “A Behavioral Model of Rational Choice.” *The Quarterly Journal of Economics* **69**(1): 99– 118.

Sofka, Wolfgang, and Christoph Grimpe. 2010. “Specialized Search and Innovation Performance – Evidence across Europe.” *R&D Management* **40**(3): 310– 23.

Spanjol, Jelena, and Charles H. Noble. 2021. “From the Editors: Introducing the New JPIM Look and Stimulating the Conversation on Accelerated Innovation.” *Journal of Product Innovation Management* **38**(2): 219– 20.

Syverson, Ben. 2020. “Rules of Brainstorming Change when Artificial Intelligence Gets Involved – Here's how.” IDEO. Accessed March, 10, 2021. https://www.ideo.com/blog/the-rules-of-brainstorming-change-when-ai-gets-involved-heres-how .

Tamkin, Alex, Miles Brundage, Jack Clark, and Deep Ganguli. 2021. “Understanding the Capabilities, Limitations, and Societal Impact of Large Language Models.” *ArXiv* **2102**: 02503.

The Guardian. 2021. “A robot wrote this entire article. Are you scared yet, human?” Accessed December 26, 2021. https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3 .

Thomke, Stefan. 1998. “Simulation, Learning and R&D Performance: Evidence from Automotive Development.” *Research Policy* **27**(1): 55– 74.

Toews, Rob. 2022. “A Wave of Billion-Dollar Language AI Startups Is Coming.” Forbes Accessed April 24, 2022.

https://www.forbes.com/sites/robtoews/2022/03/27/a-wave-of-billion-dollar-language-ai-startups-is-coming/?sh=2c27c33a2b14
.

van der Aalst, Wil. 2021. "Hybrid Intelligence: To Automate or Not to Automate, that Is the Question." *International Journal of Information Systems and Project Management* **9**(2): 5– 20.

Vaswani, Ashish, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Łukasz Kaiser, and Illia Polosukhin. 2017. "Attention Is all you Need." Advances in Neural Information Processing Systems 30, Long Beach, CA.

von Krogh, Georg. 2018. "Artificial Intelligence in Organizations: New Opportunities for Phenomenon-based Theorizing." *Academy of Management Discoveries* **4**(4): 404– 9.

Waardenburg, Lauren, Marleen Huysman, and Anastasia V. Sergeeva. 2021. "In the Land of the Blind, the One-Eyed Man Is King: Knowledge Brokerage in the Age of Learning Algorithms." *Organization Science* **33**(1): 59– 82.

Zhang, Daniel, Saurabhi Mishra, Erik Brynjolfsson, John Etchemendy, Deep Ganguli, Barbara Grosz, Terah Lyons, et al. 2021. *The AI Index 2021 Annual Report*. Stanford, CA: Stanford University.