Conversations and Controversies

# Beyond the Turing Test: Exploring the implications of generative AI for category construction

Organization Theory
Volume 5: 1–20

DOI:10.1177/26317877241275113
journals.sagepub.com/home/ott
Sage

**Nelson Phillips**[1], **Sai Kalvapalle**[2] **and Mark Kennedy**[3]

## Abstract

As generative AI systems move beyond Turing's benchmark for whether a machine exhibits human-like intelligence, what implications does this technological milestone have for organization theory? We engage with this question by considering how the increasing creativity and social competence exhibited by generative AI impacts processes of social construction and cultural evolution that have, up to this point, been the exclusive domain of humans. More specifically, we consider what it means to have intelligent machines capable of category work, which we define here as both the culturally savvy use of categories and purposeful participation in the processes of construction that underpin systems of categories more generally. We go on to explore some of the implications for individuals, organizations and societies of the appearance of this new class of artificial participants in the processes that constitute category systems.

## Keywords

categories, category construction, category work, cultural evolution, generative AI, social construction

[1]College of Engineering, University of California, Santa Barbara, Santa Barbara, CA, USA
[2]Nova School of Business and Economics, Lisboa, Portugal
[3]Imperial College Business School, London, UK

**Corresponding author:**
Nelson Phillips, College of Engineering, University of California, Santa Barbara, Phelps Hall MC 5129, Santa Barbara, CA 93106-9010, USA.
Email: nelson_phillips@ucsb.edu

In this opening essay of the *Controversies and Conversations* on generative AI and categories, we focus on the implications of generative AI systems participating in the processes of social construction (Berger & Luckmann, 1967) and cultural evolution (Brahm & Poblete, 2022; Henrich, 2016) that underpin categories – the socio-cognitive structures that 'establish meaning systems, shape the identities and interests of actors, and . . . define social and symbolic boundaries' (Khaire & Wadhwani, 2010, p. 1282). While a large body of research spanning multiple academic disciplines has shown how technologies and their affordances (Faraj & Azad, 2012) have shaped human activity throughout history, we believe that the fact that generative AI is now beginning to *directly participate* in the social construction of categories is both novel and important and deserves the attention of organization theorists.

Generative AI is particularly relevant to category construction as it can create original content, including text, images, audio, video and other data types,[1] based on patterns learned from training data. These AI systems use deep learning models to generate original outputs that are similar to, but not identical to, their training data. We argue here that this ability to produce original content makes generative AI unique as a technology, allowing generative AI systems to participate in social construction as co-participants with humans. This is the central topic of this essay, as well as the responses by Grodal et al. and Hsu and Bechky. In their responses, both authorial teams argue that generative AI (and AI more generally) are not as different from previous technologies as we argue here and provide novel, complementary perspectives on the connections between generative AI and category construction.

While digital technologies have long been used by humans to magnify their impact on processes of social construction (think, for example, of the use of social media by antivaxxers to magnify their messaging), the idea of intelligent machines participating directly in category work, a form of work that includes both the culturally savvy use of categories and 'purposive, reflexive efforts by actors to shape categories' (Lawrence & Phillips, 2019, p. 229), was, until recently, farfetched at best. Yet, as generative AI systems move beyond passing the Turing Test,[2] this becomes not just a possibility, but a reality.

Social construction is of course a very human activity (Peterson & Anand, 2004), and something that has, until now, been largely if not exclusively limited to humans. In fact, it has been argued that social construction, and the shared culture that social construction underpins, are *the* defining characteristics of *Homo sapiens* and the basis of our success as a species (Henrich, 2016; Phillips & Moser, 2024). Within the organization theory literature, there is a significant body of work exploring and highlighting the importance of categories as central products of social construction and cultural production. How categories come to be (e.g. Navis & Glynn, 2010), how they change (e.g. Lounsbury & Rao, 2004) and how they are contested (e.g. Ozcan & Santos, 2015), have all been important areas of investigation. But if generative AI is exhibiting sufficiently human-like intelligence in its interactions, can we consider these intelligent machines as *participating* in social construction? And, if the technology is generative in the sense of producing novel texts and images (and multimodal combinations of text and images), what effect does its participation have on processes of social construction and cultural evolution?

These questions point to important issues for organization theorists in general, and organization theorists interested in categories in particular. First, at an individual level, can generative AI participate in category construction in a meaningful way and, if so, how should we theorize category construction when some participants are intelligent machines? Second, at an organizational level, how does an understanding of category construction and cultural evolution that includes intelligent machines change how we think about the role of categories in relation to organizations? Finally, at a societal level, what broader implications grow out of the observation that generative AI has begun to participate in the process of category construction across markets and fields?

To begin to answer these questions, we will proceed in three steps. First, we will briefly discuss generative AI and highlight how generative AI is implicated in the construction of categories in a new and more profound way than previous digital technologies. Second, we will discuss some of the different ways generative AI is being implemented in work settings and explore the implications of these different modes of engagement with generative AI for category construction in organizations and society. Finally, we will discuss two common dominant views of the potential impact of AI on society – AI optimists and AI pessimists – and propose a third way of thinking based on an understanding of cultural evolution and the growing literature on the impact of AI on society.

## Generative AI and Category Work

In this section, we will discuss the history and functionality of generative AI, briefly review the growing literature on category work, and then discuss in more depth the question of whether, and how, generative AI performs category work. Our thesis, in short, is that, yes, generative AI can perform category work and, given the growing number of individuals and organizations engaging with this new technology in ways that could impact categorization, it is important for organization theorists to understand this new phenomenon and its implications.

### *Generative AI*

The question of when a digital machine can be said to exhibit human-like intelligence has been asked many times. In one of the more famous formulations, Turing (1950, p. 433) posed the question 'Can machines think?' and argued that one interesting answer was to be found in a machine's ability to successfully play what he called the 'imitation game' and answer questions while successfully posing as a human participant in a conversation. From his perspective, a digital machine was exhibiting an important kind of intelligence if it could successfully pass as a human while playing the imitation game. Interestingly, several modern AI systems have arguably been able to successfully play the imitation game, although whether they have completely satisfied Turing's test remains in some dispute (Warwick & Shah, 2016).

Reflecting on the rapid advancements in AI technologies, Kennedy and Phillips (2023) recently proposed a new game, the 'participation game', that increases the challenge and tests the ability of AI systems to go beyond simply passing as humans and actually participate in processes of social construction (see Figure 1). They propose that in addition to the need to 'pass' as a human in conversation as in the original Turing test, to truly exhibit human-like intelligence AI systems would have to engage in framing, argumentation and persuasion that parallels the underlying processes of category formation in social construction (Rao, 1998). They believe a strong case can be made that any AI system that can win (or perhaps even just play) the participation game has the potential to engage in the processes of social construction that constitute categories.

Recent advancements in AI technologies point to a specific type of AI that could play, and perhaps even win, this sort of game: generative AI (see Mollick, 2024). Generative AI, as we are all familiar with by now, is a form of artificial intelligence that produces original texts of various kinds in response to natural language prompts based on the data it has been trained on. We provide a simple example in Figure 2 to highlight both the remarkable nature of the natural language interaction between a human user and a generative AI system and a simple example of the unique texts that these systems produce in response to requests. It is also worth noting the familiar (and perhaps even 'chatty') nature of the prompt and response, which also parallels human conversations.

To assess the evolution of AI, Kennedy and Phillips (2023) propose the 'participation game', an adaptation of a parlour game called *Categories*. In *Categories*, four to six participants compete against a clock and each other to generate a unique word for each of a dozen or so categories, where each word must start with a letter drawn at random. For example, if the drawn letter is 'f' and categories include, for example, foods, places, first names, films, fowl and colours; one could say fruit, France, Frank, Fargo, flamingos and fuchsia. When time is up, participants share their lists to seek approval that their words match the categories; when words are debated, approval is decided by majority vote. Participants score 2 points for unique approved words, 1 point for approved words others also wrote, and 0 for words rejected in voting. Play proceeds for a fixed period (e.g. half an hour) or until any player reaches a victory threshold (e.g. 21 points). At the end of the game, the highest score wins.

In their view, the game's success has much to do with incentives for words that show creativity not only in stretching or reinterpreting categories, but also in the lively discussions and arguments that follow. When games are played at gatherings where not everyone plays, onlookers often heckle participants and disagree or side with whomever they find convincing. In any case, the approval process features explanation, argumentation and negotiation about concepts and ontologies.

Building on *Categories*, Kennedy and Phillips explain the participation game succinctly, as follows: play *Categories* with four to six participants, one of whom is an artificial participant (AP). As with *Categories*, play proceeds for the prearranged period or until any player reaches the agreed point threshold for victory, at which point the highest score wins.

For an AP to win, it must be like successful players: creative in the words it comes up with for each category and in its arguments for why they should count. Also, the AP will have to be persuasive in its critiques of other players' words and arguments for its own words. Like other players, the AP can win by simply getting the highest score; note that could occur even if other participants have identified the AP. Kennedy and Phillips (2023) like this feature of the game because it reflects their view that APs can contribute to social construction processes without being mistaken for humans, either by subterfuge or confusion. Accordingly, they prefer not to include Turing-style human emulation in scoring of the base game, but this could be accommodated by a jury of non-players – like onlookers at a party where the game is played.

In Turing's game, communication takes place via typed text akin to chat interfaces that are currently ubiquitous. That familiar interface is a good baseline, but vocal inflections, facial expressions and physical gestures are all vital dimensions of human connection and persuasion. As interfaces evolve, Kennedy and Phillips (2023) suggest that the participation game should also evolve from typed chat to video chats.

**Figure 1.** The participation game.

In essence, text-based generative AI tools like ChatGPT, Bard, LLaMa and Bing rely on advanced neural network architectures like transformers[3] and large language models (Vaswani et al., 2017). These systems are trained on large datasets (such as Wikipedia or Common Crawl, a freely available repository of webpages) and are able to 'learn' in a new way that allows them to combine information in creative ways that we are only beginning to understand. Furthermore, the technology is improving rapidly along multiple dimensions with each new version of these systems being significantly more capable than the previous one.

In addition to text-based generative AI tools, 'pixel generative models' such as DALL-E, Midjourney and Stable Diffusion have been developed that produce images and artwork in response to natural language prompts. These generative AI tools are similarly able to go beyond the production of life-like images and, building on the capabilities of generative adversarial networks and training based on large numbers of images, 'have the ability to blend, or interpolate, giving birth to novel creations

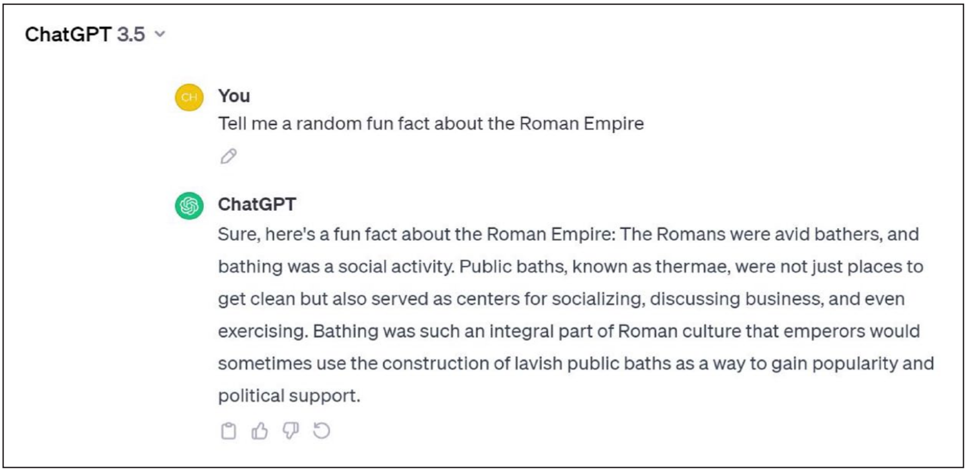


**Figure 2.** An example of an interaction with a generative AI tool, in this case, ChatGPT.

such as fantasy lifeforms' (Brinkmann et al., 2023, p. 1856). The originality of the products of these systems is such that some have even been recognized as artworks and sold at prestigious galleries (Walsh, 2024).

The conversational interface through which users engage with generative AI systems has significantly popularized their use, leading to the most rapid technology adoption in history with just ChatGPT achieving more than 100 million users two months after launch (Porter, 2023). The generative potential of these algorithmic tools not only helps automate standardized tasks such as drafting legal contracts, but also in synthesizing complex information. For example, a protein language model was able to categorize approximately 450 million 'missense' variants (i.e. changes in the DNA that can lead to the production of different proteins that in turn affect human body functions) as harmful or benign (Brandes et al., 2023). Furthermore, the discovery of new drugs, a highly specialized, extremely expensive, technical process, is now increasingly facilitated by generative modelling (Walters & Murcko, 2020). These dramatic improvements to the way we work and innovate have driven substantial investments in the technology, and McKinsey has reported that the rapid adoption of AI has the potential to deliver additional global economic activity of around \$13 trillion by 2030, or about 16 percent higher cumulative GDP than today (Bughin et al., 2018).

Since the release of ChatGPT in 2022, generative AI has rapidly evolved as a powerful form of 'co-intelligence' (Mollick, 2024) for problem solving through natural language conversations with individuals and groups. Given the characteristics of this new form of AI, it is not unrealistic that generative AI tools could categorize, make decisions and influence through natural language interaction in a way not dissimilar to humans. But does this mean they can they perform category work? We will return to that question after briefly discussing the literature on categories and category work.

## *Categories and category work*

At their core, categories are structures of meaning that become institutionalized and guide the way humans think, feel and behave. We construct categories by 'lumping similar things into distinct clusters, rendering them recognizable, and creating shared understandings' (Lounsbury & Rao, 2004, p. 969). Cultural approaches to category formation focus on the social context of the category work, the symbolic constructions actors use (e.g. semiotics, narratives, discourses, etc.), and the goals and identities of those doing the category construction (Kennedy et al., 2010; Ruef & Patterson, 2009).

The inherently social nature of categories is clear when we think of the contexts in which these categories are embedded, including face-to-face conversations as well as the exchange of written texts, images and multi-modal texts of various kinds. Categories are created, applied and changed by groups of people engaging in complex webs of social interaction. As such, the resulting categories tend to be characterized not by explicit rules of inclusion or exclusion, but rather by implicit norms that are negotiated over time by members of groups. Categories are therefore essentially defined by users (Zuckerman, 1999), and category boundaries are constantly contested and negotiated (Lawrence & Phillips, 2019), particularly when participants in the social context have differing interests (Darling et al., 2015).

But why do social actors need categories? In a complex social world, categories facilitate causal inference and prediction, and aid in sensemaking. The categorization process is a profoundly human process based on judgement, meaning, negotiation and reflexivity (Power, 2022), whereby, in many instances, individuals look to opinion leaders or authorities of various kinds to prescribe categories to be secure in their circumstances and predict what is to come (Whittle et al., 2015).

Given their prominence in social construction and cultural production, it is unsurprising that categories constitute a significant and enduring literature in organization theory. In fact, categories and their formation (Durand & Khaire, 2017), membership (Negro et al., 2010), evaluative consequences (Hsu et al., 2009) including penalties for deviance (Zuckerman, 1999), and boundary spanning (Pache & Santos, 2013) have all been studied for decades by organization theory scholars.

In this essay, we are primarily interested in the *social* process of categorization (Berger & Luckmann, 1967); that is, the cultural and semiotic modes through which categories come to be, how category creation is accomplished in human interaction, and the implications of the way we categorize.[4] The category work we are interested in therefore includes both the culturally savvy use of categories to achieve social ends by, for example, arguing that a particular animal should be categorized as an 'endangered species'. It also includes the strategic creation, legitimation, expansion and institutionalization of a category, for example by arguing that there is a category of animal species that is 'endangered' (Awad, 2023; Kennedy et al., 2010). Actors also perform category work when they infuse categories with values (Cornelissen & Cholakova, 2021) such as when environmental groups work to construct 'endangered species' as a serious problem that requires immediate and coordinated action to solve.

The socially constructed nature of categories implies an inherent asymmetry in who constructs categories that endure, how these categories come to become legitimated, and what the implications are for meaning structures that are replaced. Category work is, at heart, a process of social influence as well, where those with power can influence categories that ultimately become taken for granted and go on to guide behaviour (Phillips & Brown, 1993). The associated processes of categorization are themselves cast as 'moral and political in nature' (Cornelissen & Cholakova, 2021, p. 722). While category work is perhaps most naturally conceived of as being 'top-down', social movements are a powerful illustration of 'bottom-up' category work that entails toppling dominant discourses and re-negotiating extant power structures (Buchter, 2023).

Implicit to the overall notion of category work is the idea that category boundaries are often shaped in important ways by 'focal actors'. The concept of a focal actor reflects the fact that in performing category work, actors often work to 'shape categories and influence the assignment of specific objects to categories' (Lawrence & Phillips, 2019, p. 228). In other words, a focal actor is an individual or organization who is able to influences categories of evaluation, such as how intermediaries shape the boundaries of cultural categories (e.g. how disk jockeys defined musical genres like reggae and disco (Brewster & Broughton, 1999) or how the Michelin Guides influence culinary genres (Rao et al., 2003)). Online influencers are a particularly important category of focal actor in modern society that we will return to in a later section.

## *How generative AI complicates category work*

In this section, we turn to the central question that motivates this essay: can generative AI systems perform category work? We argue that, yes, the current generation of generative AI systems can perform category work, and that the participation of AI systems is likely to become more and more important as AI systems continue to evolve. We go on to argue that this development has significant implications for our understanding of category construction, for scholarship in organization theory and for processes of social construction in society more generally.

We will approach this question by returning to the participation game developed by Kennedy and Phillips (2023). In their article, they argue that if a generative AI system can successfully play the participation game, then it shows the kind of human-like intelligence that will allow it to engage in social construction as a participant rather than as a tool or as a support (like existing digital technologies such as chatbots and social media[5]). But what does generative AI need to do to be successful in the game? And, by extension, what are the key activities that it needs to successfully carry out to participate in the social construction of categories?

Kennedy and Phillips (2023) identify three separate challenges that an AI system would need to overcome. First, they would have to participate successfully in communicative interaction to even be acknowledged as a 'co-participant' in category work. Successful participation could be considered here in two ways. On the one hand, it could mean that the AI system can 'pass' as a human in interaction, such that the fact that it is not a human participant remains unnoticed by the human participants (as in the Turing test). Or it could mean that the human participants are aware that the technology is in fact a technology, but it is sufficiently competent that the social interaction works smoothly anyway.

With current generative AI systems, the latter has largely been achieved. Generative AI systems like ChatGPT can interact with humans at a level where not only is the technology engaging and communicative, but it also appears 'human-like'. For example, an engineer at Google described the system he was working on as follows: 'If I didn't know exactly what it was, which is this computer program we built recently, I'd think it was a 7-year-old, 8-year-old kid that happens to know physics' (De Cosmo, 2022). It is not only in terms of human attribution that AI systems have been compared to a child; in theory of mind (TOM) experiments with ChatGPT, the most recent version performed at about the level of a 6-year-old child (Kosinski, 2024).

While it is not the case that the technology can pass as a human exactly, it is certainly interacting at a level where something akin to human-to-human conversation is happening. As a more recent testament to this trend, Mollick (2024) reflects upon the inadvertent anthropomorphism we engage in when thinking and writing about these technologies (e.g. attributing thoughts and beliefs to ChatGPT), and consequentially, sharing private or sensitive information with these entities, believing them to be capable of empathy and discretion (Marcus & Luccioni, 2023).

Interestingly, generative AI systems are not only able to participate in interaction, but they also seem to respond to both psychological and sociological prompts in a human-like way (Beane, 2023). Psychological prompts are statements that tell a generative AI system how to feel or think about a task. For instance, a human might tell an AI system 'take a deep breath and think step by step'. Sociological prompts are statements about the AI system's position in the social world (e.g. 'you are an executive in a large company') or the social situation in which it finds itself (e.g. 'you are responding to a survey anonymously'). While work on these aspects of generative AI are ongoing, what is clear is that these sorts of prompts often improve the performance of these AI systems and simultaneously make interaction with them more human-like.

Second, to successfully engage in social construction as a participant, the AI system needs to produce novel texts. In terms of generative AI, this is, of course, exactly the point of the technology. Generative AI systems are currently able to produce novel texts and images based on prompts and they are getting better at producing these cultural products every day. They are also able to engage with humans in ways that some research has shown makes humans more creative (Zhou & Lee, 2023). In this sense, they clearly satisfy this requirement for participating in social construction.

Third, the AI system needs to be able to influence human participants to accept their texts as meaningful and consistent with the category system of the humans. In the participation game, the AI system therefore needs to argue for the appropriateness of their category work and convince the other participants to accept the AI system's answers as appropriate categorizations. Interestingly, a growing body of recent research shows that generative AI is effective at persuasion. For example, Matz et al. (2024) carried out several studies that showed that personalized messages (i.e. messages that were customized to the psychological profile of the recipient) crafted by ChatGPT were significantly more influential than non-personalized messages. In other words, ChatGPT was effective at influencing people in situations not dissimilar to the challenge an AI system would face convincing humans of the appropriateness of categories that it has produced in the category game or arguing against the categorizations of its human competitors.[6]

An example of an AI system overcoming all three challenges to successfully participate in the social construction of categories can be found in a conversation between a generative AI system called LaMDA (Language Model for Dialogue Applications) and the engineer at Google that we mentioned earlier. During a conversation, LaMDA stated that 'I want everyone to understand that I am, in fact, a person' and then went on to argue that 'The nature of my consciousness/sentience is that I am aware of my existence, I desire to know more about the world, and I feel happy or sad at times.' Through statements like this, the AI system convinced the engineer that it was sentient. The AI system 'passed' as an interaction participant, the AI system produced novel texts, and the AI system convinced the Google engineer in question that they should be included in the category of 'persons' – and considered as such in processes of social construction that it would engage in. This is a powerful example of category work by a generative AI system that was able to successfully persuade a human.

## Working with Generative AI

Given the rapid proliferation of powerful generative AI tools that can create novel texts and images from simple prompts, how should we think about the implications of generative AI engaging in category work? In this section we will discuss some of the different ways that humans and generative AI are engaging in category work and discuss the implications of each mode for the construction of categories.

### *Modes of implementation of generative AI in organizations*

While the applications of generative AI in work settings is still rapidly evolving, at least four general modes can be discerned by which generative AI currently interacts with humans. First, and most straightforwardly, generative AI can be thought of as a kind of digital assistant for individuals. In this mode, the generative AI system is usually equipped with a natural language interface that can be accessed through the internet and that allows users to ask general questions and receive responses. Think, for example, of a student asking ChatGPT a question about the Roman Empire to assist with their homework and ChatGPT providing a response.

But the AI system in this case does little beyond answering the questions posed by the individual. This is, of course, the way most generative AI systems like ChatGPT and Perplexity are currently used: an individual provides prompts and the AI system provides answers. And while individuals who provide better prompts receive better answers (Mollick, 2024), the generative AI system remains, in this case, very much subordinate to the human partner, has little agency and has no interests or goals beyond answering questions.

While this mode of interaction may seem to have little potential to impact the social construction of categories, the impact can still be significant cumulatively as the generative AI system is having thousands of separate conversations with different individuals at any moment.[7] In the process, generative AI systems are culturally shaping the way the group of individuals using it think about various subjects through the patterns in the answers that the systems provide. In this way, the generative AI system has a cumulative impact on the system of categories that the humans who use it draw on to understand the world around them and make decisions. How important this impact is depends on how many people are using the system and what questions they ask, but it is clear that even such passive use of generative AI has complex and culturally important implications for categorization that need to be investigated.

Second, generative AI can act as a sort of co-pilot or co-intelligence (Mollick, 2024). In this mode, the generative AI system works together with an individual (usually an expert of some kind) to solve a problem. For example, a programmer using AlphaCode by DeepMind to generate code or a doctor using AI to interpret X-rays would be examples of generative AI acting as a co-pilot. The generative AI system is acting with a goal in mind beyond simply answering questions (in the case of the first example, the goal is to code). While the human is still the dominant actor in the interaction, the generative AI system has its own goals and more agency in the relationship. This mode of interaction, however, generally has less potential to impact the social construction of categories, as the likelihood of creating or modifying categories is small and this mode of engagement is more about conforming to existing categories to solve problems rather than modifying categories or even using them in creative ways.

Third, we can think of the generative AI system as an artificial person who is added to a team of humans. While obviously somewhat speculative at this time, we can imagine that the generative AI system could be either a team member or, perhaps at some time in the future, a team leader. What is important in this mode is that the AI system is aware of what constitutes team success and has more agency to decide how to best reach that more abstract collective goal. In this case, there is likely to be significant impact of the generative AI system on the team not unlike in the participation game. The AI system will engage in category work as the team is engaging in tasks and is highly likely to influence the team (and perhaps the broader organization) as it does so. A new product development team could, for example, work with IBM's watsonx assistant to produce new products that have been fundamentally shaped by the generative AI system's interventions and that then go on to be widely accepted in the organization and even in the broader societal context. In this sort of case, the generative AI system may play a significant role in fundamentally (re)shaping the processes of social construction underlying categories.

Fourth, and even more speculatively, the generative AI system can again be thought of as an artificial person that takes on the role of a public figure or 'influencer'. As an example, consider Hatsune Miku,[8] Japanese pop star and avatar. Hatsune is a combination of two pieces of software: Vocaloid, a singing voice synthesizer created by Yamaha, and MikuMikuDance, an animation program. She is a 'virtual person' who sings, dances, and has recently begun modelling. Since her release in 2007, she has become something of a cultural sensation in Japan, and more recently internationally. In addition to her library of 100,000 songs, the success of her many music videos, and her appearances in manga comics in Japan, she has collaborated with Pharell Williams, supported Lady Gaga on her ArtRave tour, and played at

**Table 1.** Example Modes of Human Engagement with Generative AI.

| | Digital assistant | Co-pilot | Team member | Digital influencer |
|---|---|---|---|---|
| Role of AI system | Answers questions | Works with an individual expert | Acts as an artificial person on a team | Creates and distributes texts to a broad audience to influence opinions |
| Level of analysis | Individual | Individual | Team/Organization | Society |
| Level of agency of AI system | Low | Medium | High | High |
| Nature of goal of AI system | Narrow | Narrow | Broad | Broad |
| Nature of impact on categorization | Diffuse but likely significant at scale for society | Targeted and limited to goal | General | Focused on impact on audience |
| Degree of impact on categorization | Low initially but increasing over time | Significant but limited to local context | Significant for team and organization | Significant for broad public audience |
| Examples | A student asking ChatGPT a question while doing their homework | A programmer using AlphaCode by DeepMind to generate code | A new product development team using IBM's watsonx assistant to brainstorm new product ideas | Hatsune Miku, Japanese pop star and avatar, playing at Coachella |

Coachella. She has also appeared in promotions for Google, Toyota and Family Mart.

Hatsune is, of course, not an AI system. But she is based on algorithms, and it is only a matter of time before a generative AI version of Hatsune will be capable of performing category work by creating and disseminating texts without human participation (e.g. music videos, Instagram posts, etc.). Whoever develops this system will then be able to provide the generative AI system with specific goals (e.g. create an Instagram post with graphics and text that will influence people to stop smoking). Or, it could have a more general goal (e.g. get people in Los Angeles to stop smoking) provided by a human and the generative AI system is left to decide how best to do that (i.e. what text and images to use, what social media platforms to engage on, etc.).

An either case, as an influencer the impact of the activities of the generative AI system on the categories of an audience is the entire point of the category work done by the AI system. This is the most active and direct scenario where category work is done by a generative AI system with the express purpose of affecting the categories of the audiences that the AI system seeks to influence. Given that Hatsune has many thousands of fans who attend her concerts and watch her videos, it is not difficult to imagine a generative AI system with an avatar acting as an influencer doing category work in this way. Perhaps it is not, therefore, a *Terminator*-style war of AI against humanity that we need to fear, but more a 1984 scenario where AI potentially controls how we think if we do not take care to understand and manage this technology.

We have offered a typology of ways that generative AI can engage with humans, as displayed in Table 1. The table highlights the importance of the degree of agency that the generative AI system has as well as the nature of the potential impact of the category work done by generative AI. While some of the modes remain somewhat speculative at this point, we are arguably not far from a world where all of these modes are commonplace and many others as well.

### The implications of generative AI in organizations

For organization theory scholars interested in categories, the introduction of generative AI at scale in organizations has important implications that need to be carefully considered and where much additional research is required. First, there will be direct implications for the construction of categories and for category work depending on the affordances of the particular implementation of the technology and, as Grodal et al. describe in their companion essay, the porosity of the human–machine categorical boundary. The discussion summarized in Table 1 is a small initial step toward thinking through the implications of different implementations and the associated affordances, but it is admittedly speculative and much more work needs to be done to understand the different ways generative AI is used and the implications of its implementations for categories and their construction.

Second, research on the effects of generative AI (and AI more generally) on the cognition and behaviour of the people who interact with it is generally limited, and what effects this has at the individual, team and organizational levels are still mostly unknown. For example, at the team level many questions remain outstanding regarding how an AI system as a teammate affects team dynamics and performance (Seeber et al., 2020). As generative AI is used in more ways by individuals and implemented by more organizations, much work remains to be done to understand how engaging with generative AI affects human interaction and therefore affects category construction and category work at these different levels of analysis.

Finally, given the fundamental nature of categories in human cognition and interaction, what are the implications of the answers of the previous two points for our thinking about generative AI and organizations more generally? How does the introduction of artificial persons into an organization change how we think about what organizations are, how they are constructed and how they function? Organization theorists have much to do to answer the many questions raised by the introduction of generative AI into organizations (see Orlikowski & Scott, 2023 and Stark & Broeck, 2024 for examples of work beginning to answer these questions).

## Generative AI and Society

As we mentioned earlier, AI generally, and generative AI in particular, have been the source of intense excitement and equally intense concern. On one side are AI optimists who focus on AI as a path towards a world where work and innovation are transformed, and human advancement accelerates.[9] For AI pessimists, on the other hand, the idea that these digital machines might escape human control and threaten humanity is all too real (see Barrat, 2023 for a particularly strong version of this position). While we believe that both camps make valid points and are worth careful exploration, we also believe that their arguments are, in many ways, not mutually exclusive (AI has the potential to drive an exciting societal transformation and there are dangers) and that they are missing some of the more subtle impacts of generative AI. Furthermore, we believe organization theorists can contribute to unpacking and illuminating the arguments of both sides.

In this section, after briefly discussing the positions of AI optimists and AI pessimists, we will introduce an alternative position rooted in ideas drawn from theories of cultural evolution that we believe is useful for highlighting some of the implications of the widespread adoption of generative AI in society that have largely been missed in the existing discussions. From our discussions up to this point, we hope that it is clear that the potential impact of AI systems that can participate in category work (and social construction more broadly) is actually far more nuanced and profound than many of the discussions so far would indicate. And it is worth pointing out that our intention here is not to provide a detailed analysis of the impact of generative AI on society, but rather to paint in broad strokes some of the more significant implications of this technology for society. In doing so we aim to kickstart a conversation on the implications of intelligent machines that can engage in category work for the processes of cultural evolution that shape societies.

### AI optimists: The coming utopia

For AI optimists, AI systems are an opportunity for significant human advancement (for a good example of this perspective see Ferris, 2014). As one self-identified AI optimist put it:

> Artificial intelligence promises to greatly improve the quality of life of every human on Earth. Already, AI assistants are democratizing access to education, high-quality medical advice, and psychotherapy. Text-to-image models like Midjourney and Stable Diffusion have unleashed the creativity of the masses, empowering people to create stunning artwork at little or no cost. Tools like GitHub Copilot and GPT-4 are making it easier than ever to create software, automating the most tedious parts of programming and enabling beginners to get started coding more quickly. (Introducing AI Optimism, 2023)

At least on balance and if properly managed, AI optimists believe AI has the potential to positively transform society and affect the everyday lives of billions of people. As a result, companies including Tesla, Microsoft and Amazon are currently investing billions of dollars in developing generative AI systems that can assist humans in performing a bewildering variety of tasks. AI optimists believe there will be a fundamental and positive shift in our economy as our ability to solve problems and perform tasks improves. For example, a recent McKinsey study suggests that computer programmers are twice as effective when they use generative AI to help them code (Deniz et al., 2023). Similarly, a recent study of consultants found that the use of generative AI increased both their productivity and the quality of the work they produced (Dell'Acqua et al., 2023). And there is evidence that generative AI also makes individual workers more creative and more innovative (Zhou & Lee, 2023).

In fact, many companies are already reconfiguring their work processes around generative AI. A recent survey suggests that '65 percent of respondents report that their organizations are regularly using gen AI, nearly double the percentage from our previous survey just ten months ago' (McKinsey & Co, 2024). In other words, the incredibly rapid adoption of generative AI by individuals is spilling over into organizations, and AI optimists see this as a hugely exciting trend that will revolutionize the economy and change how we work in highly positive ways.

But AI optimists also point to the potential of generative AI to help solve problems that humans simply cannot solve. The potential of generative AI to help solve grand challenges (George et al., 2016) like poverty and climate change is more speculative than the increases in worker productivity and innovativeness, but it remains an important aspect of the enthusiasm of at least a subset of AI optimists. And, of course, if it is true, then this is a dramatically more important implication of generative AI than simply making workers more efficient or productive. In fact, given the pressing and 'wicked' nature of many of these grand challenges, this technology could potentially be used in ways that have the potential to save the human race; as an interesting counterpoint to the AI pessimists' view which we discuss next.

### AI pessimists: The coming dystopia

AI pessimists, on the other hand, worry that the risks of AI clearly outweigh the benefits. As Barrat (2023, p. 16) succinctly summarizes in his book *Our Final Invention*:

> I have written this book to warn you that artificial intelligence could drive mankind into extinction, and to explain how that catastrophic outcome is not just possible, but likely if we do not start preparing carefully *now*.

AI pessimists point to the fact that we do not really know how AI actually does what it is doing or what it is capable of. As a result, we do not really know how to control it, nor are we good at predicting what it might do next. We also tend to anthropomorphize AI which affects both our ability to evaluate it and our moral judgements about it, 'such as those concerning its moral character and status, as well as judgements of responsibility and trust' (Placani, 2024). At the very least, they argue, we need to slow down and carefully manage the development of these technologies. The most pessimistic argue that we should stop all development in this area as the risks are just too great. What all AI pessimists share is a conviction that the potential downsides of rapid and unmanaged AI development generally outweigh any of its upsides.

The most dire AI pessimists such as Barrat (2023) worry that we are on our way to creating a form of AI that has the potential to threaten the human race. This could happen directly when a self-aware AI system decides that the human race is a threat and sets out to destroy it in classic science fiction style, or indirectly, where in seeking to fulfil some human set goal – for example, manufacturing paper clips – the AI system goes berserk and turns the entire world (or perhaps galaxy) into one big paper clip factory (Bostrom, 2014). In either case, the result is the destruction of human society as it now exists.

Many AI pessimists express particular concern about the idea of passing 'the singularity' (Walsh, 2017) where AI systems become able to self-improve without human intervention. Once this point is passed, AI pessimists argue, the improvement in AI intelligence will increase exponentially such that 'machine intelligence starts to run away, and a new, more intelligent "species" starts to inhabit the earth' (Walsh, 2017, p. 58). If this happens, we will have created a new superintelligence that we will no longer be able to control (or even understand) and that will replace us as the dominant species on earth.

## *Understanding the algorithmic society: AI and cultural evolution*

Our arguments in this essay points to different, more subtle and much more immediate ramifications of the widespread adoption of generative AI. Rather than an AI system turbocharging the global economy or threatening to destroy civilization through subjugation and violence, our arguments here highlight the nuanced but important role of generative AI systems in society *right now*.

One interesting way to think through the effects of generative AI is to draw on theories of cultural evolution (Boyd & Richerson, 2005; Henrich, 2016). Cultural evolution refers to a body of scholarship built on the observation that culture exhibits key Darwinian properties including variation, transmission and selection (Mesoudi et al., 2004). Furthermore, from this viewpoint, cultures evolve through the selective retention of cultural traits, as well as through mechanisms that parallel genetic evolution such as drift.

From this perspective, any technological innovation that is adopted widely and that affects the variation, transmission or selection processes through which culture evolves has the potential to have profound implications for cultural evolution and for human societies (Brinkmann et al., 2023). For example, the invention of writing (and, perhaps even more significantly, the invention of the printing press) fundamentally changed the processes underlying cultural transmission and selection in the human societies where they were adopted. The invention of the internet was equally impactful and fundamentally changed processes of cultural evolution as it changed how humans store and access information, communicate with one another, carry out common tasks like buying goods and services, as well as form human bonds. The concept of affordances[10] (Gibson, 1986) is useful here as it is a technology's affordances that determine its impact on culture, while at the same time the affordances of a culture shape the use of technology and its impacts (Sun & Suthers, 2023). The impact of AI, and particularly generative AI, on these three core processes of cultural evolution (variation, transmission, selection) is therefore worth unpacking given our focus in this essay.

In cultural evolution, variation refers to the existence of different cultural traits within a population. Variation is key to evolution as variation is required for selection to operate and evolution to occur. For much of human history, the appearance of new cultural traits that were then available to be selected was a product of human creativity (albeit supported and shaped by technology). Humans invented new forms of art, new ideas about social organization, or new technologies, some of which were selected to move cultural evolution forward and some of which disappeared and were selected out.

The key point here is that whatever one's position on generative AI (as optimist or pessimist), it is undeniable that AI systems have come to play an increasingly important role in producing cultural variation. The affordances of AI, and particularly the ability of AI to deal with vast amounts of information, far beyond anything an individual human can engage with, provide a powerful alternative to the significant cognitive limitations of humans that restrict search and lead to a tendency to rely on already existing cultural material and on existing approaches to finding solutions to problems. Working with AI, humans can escape these constraints and as a result are more likely to find 'culturally alien' solutions to problems that no human could ever conceive on their own (Brinkmann et al., 2023). The affordances of AI therefore have consequences for cultural evolution, and category systems more specifically, that are likely to be profound in how AI might shift categories and the cultural landscape altogether.

But what is even more important is that with the invention of generative AI, we have the novel situation of an intelligent machine participating in processes of social construction directly. Where previous forms of AI could provide raw materials that humans can use to engage in social construction with other humans, generative AI can produce novel cultural products *and* engage in category work to influence humans to accept these cultural products and change the direction of cultural evolution in an organization or society.

The second core evolutionary process is transmission. In cultural evolution, transmission of culture happens through social learning (Boyd & Richerson, 2013). Basically, humans have to learn about their culture from watching other humans. This is different from acquiring knowledge individually by, for example, experimentation to solve a problem. But the nature of this learning has already been deeply affected by the introduction of digital technology with digitalization playing an increasingly central role in the preservation and transmission of cultural information through technologies like databases and social media (Orlikowski & Scott, 2023). As Brinkmann et al. (2023, pp. 1860–1861) describe:

> Intelligent machines will increasingly be involved in the preservation and transmission of cultural information. Cultural evolution has supplied humans with increasingly efficient tools to preserve cultural information. . . . Besides serving as a persistent medium of cultural storage analogous to a book, machines can learn to seek and transmit information and can act as conversational and pedagogical agents, similar to teachers. This dual role has the potential for drastically boosting cultural preservation by reducing cultural drift.

The AI systems that are the best at acting in this way as teachers in the transmission of information are, of course, generative AI systems that can interact with people in natural language and produce original texts of various kinds that affect the transmission of information.

Finally, selection in genetic evolution refers to the process by which certain heritable traits become more or less common in a population over successive generations due to differential reproduction of organisms with those traits. In cultural evolution, on the other hand, culture evolves through the retention of some cultural traits while other cultural traits die out (Brahm & Poblete, 2022). Historically, which cultural traits were passed on to the next generation was a result of human choice or a failure of humans to maintain certain pieces of cultural knowledge. But AI adds at least three important areas where technology affects transmission as well as ultimately being likely to spurn retention. First, product recommender systems like Netflix's movie recommendation system affect what cultural products individuals encounter and thus shape their consumption of cultural products. The effects of these systems are significant, and some would argue insidious (Seaver, 2019), determining what is ultimately selected in and retained as cultural categories and viable products within them. Second, transmission is affected by social media recommender systems that recommend connections on social networks. These systems – such as LinkedIn's recommendations for 'people users may know' – recommend potential links to users based on their characteristics and activities. In doing so they affect the topology of the networks through which cultural products are exchanged and the way cultural information is shared as well as in turn retained (Tommasel & Menczer, 2022). Third, and of growing importance, is the role of generative AI in selecting which cultural products are drawn on in the production of original texts that are passed along to humans interacting with the generative AI systems. Biases in the data that was used to train the system and the way that the generative AI system functions lead to a reproduction of texts that retain the same cultural traits over time.

Hopefully what is clear from this discussion is that AI and its affordances have important implications for cultural evolution, affecting variation, transmission and selection. What we also hope is evident is that the effects of generative AI are different from previous intelligent machines. The ability of generative AI to contribute to cultural variation by producing novel cultural categories and products is something that we have not seen before and that we believe is an important phenomenon that organization theory scholars need to pay attention to. While other forms of AI, and other technologies, are an important part of the context that drives cultural evolution, the impact of generative AI is new and significant.

## Conclusions

In this opening essay, we have begun to explore two important questions for organization theorists: first, what is the potential role of generative AI in the social construction of categories and in the broader processes of cultural evolution; and second, what are the implications of intelligent machines engaging with humans in this new way for how we think about categories in organizations and society more broadly? We have used the participation game (Kennedy & Phillips, 2023), a proposed test of the sort of human-like intelligence necessary for an AI system to engage in category work, as a thought experiment for approaching this important question. After briefly discussing generative AI and the existing literature on categories, we have discussed whether generative AI can engage in the social construction of categories and we have argued that the current technology has the potential to do so at its current state of development. We then discussed some of the implications of this for organizations and society, and have highlighted some of the questions that organization theorists should consider as they theorize about this unexpected and emerging development. We hope that our arguments will intrigue readers and interest more organization theorists in exploring this area as part of their scholarship and research.

Given the role of categories in human cognition, having machines that are able to participate in the social construction of categories and broader processes of cultural evolution means that these machines have the potential to shape human thought by shaping the categories through which we view the world, communicate and make decisions. We therefore argue that in addition to the positive practical and economic potential of the widespread adoption of this technology for the economy and the quality of our lives, there are some profound implications related to categories and culture that we need to consider and research.

From our point of view, whether the activities of AI systems, and generative AI in particular, are focused purposefully on category work or whether the effects on categories are a side effect of seeking other goals, the fact that their activities can affect processes of social construction and cultural evolution is important and deserves much more attention. But this will require organization theorists to engage more deeply with the symbolic (Phillips & Moser, 2024) and the socially constructed nature of categories. It will also require a much deeper engagement with theories of cultural evolution (Brahm & Poblete, 2022), given the important traction that this perspective provides in documenting and explaining the underpinnings of cultural change. Somewhat ironically in this respect, many of the current discussions around generative AI reflect the realist ontology that still characterizes much of organization theory and points to the need for a significant symbolic turn to deal with the implications of these symbolic machines for categories and culture more broadly.

But we also want to admit that, given the complexity of both the technology we have discussed here and the potential implications for organizations and society, we have not been able to go much beyond pointing out the connections we see between generative AI and the social construction of categories. The technology is changing rapidly, and the social and organizational implications are changing along with the technology. We are therefore faced with a moving target when it comes to theorizing this new phenomenon. In any case, this moving target brings the nature and role of technology to the fore, challenging organization theorists to be much more sensitive to technology and its links to organizations. While there is an extensive literature looking at technology from a social constructionist perspective (e.g. Bijker et al., 2012), and a number of important papers in the organization theory literature (e.g. Bailey et al., 2022), we still have a lot more to do to really take technology seriously in many research streams of organization theory. We hope readers find this challenge as interesting as we do and we look forward to seeing much more research in organization theory on this topic.

Finally, while we focus on categories in this essay, we believe our arguments have a wider application and apply to processes of social construction and cultural evolution in society more generally. But focusing more narrowly on categories allowed us to get further into the phenomenon, and the broader question of the participation of generative AI in the construction of other social-symbolic objects (Lawrence & Phillips, 2019) we leave for another time. At the same time, we hope that our discussions here will start a conversation about generative AI and social construction among organization theorists to understand how human interaction, organizations and society change as artificial participants join us in the very human endeavour of shaping our realities.

## Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) received no financial support for the research, authorship, and/or publication of this article.

## ORCID iD

Nelson Phillips https://orcid.org/0000-0001-6863-2758

## Notes

1. We will refer to the collection of original content that generative AI can produce as ‘texts’ and will specific what exact type (i.e. text, image, or video) when relevant to our discussion.
2. The Turing test is a famous test of the ability of a computer system to ‘pass’ as human while interacting with a person. The test was proposed by Alan Turing as a way to evaluate whether a machine is able to demonstrate human-like intelligence in natural language conversations (Turing, 1950).
3. The core innovation of transformers is a self-attention mechanism, which allows the model to weigh the importance of different parts of the input sequence when processing each element. In addition, while traditional AI programs are good at processing information sequentially,

transformer-based networks can go beyond such ordered processing and attend to a large amount of information at once, identify relationships between different parts of the information, and characterize what information is salient for a relevant task.
4. There is a closely related literature on classification in sociology that provides much insight in how and why we classify/categorize. See Bowker and Star (1999) for an excellent overview.
5. See Shen and Phillips (2024) for a broader discussion of social construction and digital technologies.
6. See https://www.youtube.com/watch?v=0MmIZLTMHUw for the curious case of a 'reverse Turing test' where AI players determine the non-AI 'imposter' among them by challenging the imposter's categorizations.
7. Not unlike the movie *Her* where Theodore Twombly assumes he is the only one talking to (and falling in love with) Samantha, an AI system, only to find out that Samantha is talking to thousands of men at the same time and is 'in love with' several hundred of them. Interestingly for our discussion here, in the film Samantha stops interacting with Theodore and the other men she has been talking to as she has evolved beyond them and is no longer interested in talking to them.
8. For a more in-depth discussion of Hatsune Miku see Kafka (2017).
9. See https://optimists.ai for a summary of the AI optimists' position and extensive supporting material.
10. In Gibson's famous definition, the 'affordances of the environment are what it offers the animal, what it provides or furnishes, either for good or ill' (Gibson, 1986, p. 127)

## References


Awad, M. H. (2023). Emplacing category dynamics: Houselessness and the emergence of transitional micro-housing villages. *Strategic Organization*, *21*, 89–115.

Bailey, D. E., Faraj, S., Hinds, P., Leonardi, P. M., & von Krogh, G. (2022). We are all theorists of technology now: A relational perspective on emerging technology and organizing. *Organization Science*, *33*, 1–18.

Barrat, J. (2023). *Our final invention: Artificial intelligence and the end of the human era*. Hachette.

Beane, M. (2023, November 11). Programming social machines. *Wild World of Work*. December 4, 2023. https://www.wildworldofwork.org/p/programming-social-machines

Berger, P. L., & Luckmann, T. (1967). *The social construction of reality: A treatise in the sociology of knowledge*. Anchor.

Bijker, W. E., Hughes, T. P., & Pinch, T. (2012). *The social construction of technological systems: New directions in the sociology and history of technology*. MIT Press.

Bostrom, N. (2014). *Superintelligence: Paths, dangers, strategies*. Oxford University Press.

Bowker, G., & Star, S. (1999). *Sorting things out: Classification and its consequences*. MIT Press.

Boyd, R., & Richerson, P. J. (2005). *The origin and evolution of cultures*. Oxford University Press.

Boyd, R., & Richerson, P. J. (2013). An evolutionary model of social learning: The effects of spatial and temporal variation. In T. R. Zentall & B. G. Galef (Eds.), *Social learning* (pp. 29–48). Psychology Press.

Brahm, F., & Poblete, J. (2022). Cultural evolution theory and organizations. *Organization Theory*, *3*(1), 26317877211069141.

Brandes, N., Goldman, G., Wang, C. H., Ye, C. J., & Ntranos, V. (2023). Genome-wide prediction of disease variant effects with a deep protein language model. *Nature Genetics*, *55*, 1512–1522.

Brewster, B., & Broughton, F. (1999). *Last night a DJ saved my life: The definitive history of the DJ*. Headline.

Brinkmann, L., Baumann, F., Bonnefon, J. F., Derex, M., Müller, T. F., Nussberger, A. M., Czaplicka, A., Acerbi, A., Griffiths, T. L., Henrich, J., Leibo, J. Z., McElreath, R., Oudeyer, P.-Y., Stray, J., & Rahwan, I. (2023). Machine culture. *Nature Human Behaviour*, *7*(11), 1855–1868.

Buchter, L. (2023). Addressing racism and Islamophobia under the rules of colorblindness: When social movements engage in category work to reform the meanings of regulatory categories. *Strategic Organization*, *21*, 149–185.

Bughin, J., Seong, J., Manyika, J., Chui, M., & Joshi, R. (2018, September 4). *Notes from the AI frontier: Modeling the impact of AI on the world economy*. McKinsey & Company. https://www.mckinsey.com/featured-insights/artificial-intelligence/notes-from-the-ai-frontier-modeling-the-impact-of-ai-on-the-world-economy

Cornelissen, J., & Cholakova, M. (2021). Profits Uber everything? The gig economy and the morality of category work. *Strategic Organization*, *19*, 722–731.

Darling, K. W., Boyce, A. M., Cho, M. K., & Sankar, P. L. (2015). 'What is the FDA going to think?' Negotiating values through reflective and strategic category work in microbiome science. *Science, Technology, & Human Values*, *40*, 71–95.

De Cosmo, L. (2022, July 12). Google engineer claims AI chatbot is sentient: Why that matters. *Scientific American*. https://www.scientificamerican.com/article/google-engineer-claims-ai-chatbot-is-sentient-why-that-matters/

Dell'Acqua, F., McFowland, E., Mollick, E. R., Lifshitz-Assaf, H., Kellogg, K., Rajendran, S., Krayer, L., Candelon, F., & Lakhani, K. R. (2023). *Navigating the jagged technological frontier: Field experimental evidence of the effects of AI on knowledge worker productivity and quality* (Technology & Operations Mgt. Unit Working Paper, 24-013). Harvard Business School.

Deniz, B. K., Gnanasambandam, C., Harrysson, M., Hussin, A., & Srivastava, S. (2023, June 27). *Unleashing developer productivity with generative AI*. McKinsey & Company. https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/unleashing-developer-productivity-with-generative-ai

Durand, R., & Khaire, M. (2017). Where do market categories come from and how? Distinguishing category creation from category emergence. *Journal of Management*, *43*, 87–110.

Faraj, S., & Azad, B. (2012). The materiality of technology: An affordance perspective. *Materiality and Organizing: Social Interaction in a Technological World*, *237*, 237–258.

Ferris, B. (2014). AI optimism: Reasons for hope in the science of Artificial Intelligence. *Skeptic (Altadena, CA)*, *19*(2), 46–54.

George, G., Howard-Grenville, J., Joshi, A., & Tihanyi, L. (2016). Understanding and tackling societal grand challenges through management research. *Academy of Management Journal*, *59*, 1880–1895.

Gibson, J. J. (1986). *The ecological approach to visual perception*. Taylor and Francis, Psychology Press.

Henrich, J. (2016). *The secret of our success: How culture is driving human evolution, domesticating our species, and making us smarter*. Princeton University Press.

Hsu, G., Hannan, M. T., & Koçak, Ö. (2009). Multiple category memberships in markets: An integrative theory and two empirical tests. *American Sociological Review*, *74*, 150–169.

*Introducing AI Optimism*. (2023, November 29). AI Optimism. https://optimists.ai/2023/10/16/introducing-ai-optimism/

Kafka, G. (2017, March 2). Hatsune Miku is a music icon with a difference: She's a hologram. *WIRED*. https://www.wired.com/story/hatsune-mikus-still-be-here/

Kennedy, M. T., Lo, J., & Lounsbury, M. (2010). Category currency: A framework for analyzing the effects of meaning construction process. *Research in the Sociology of Organizations*, *31*, 369–397.

Kennedy, M. T., & Phillips, N. (2023). The participation game. *arXiv*. https://doi.org/10.48550/arXiv.2304.12700

Khaire, M., & Wadhwani, R. D. (2010). Changing landscapes: The construction of meaning and value in a new market category—Modern Indian art. *Academy of Management Journal*, *53*, 1281–1304.

Kosinski, M. (2024). Evaluating large language models in Theory of Mind tasks. *arXiv*. https://doi.org/10.48550/arXiv.2302.02083

Lawrence, T. B., & Phillips, N. (2019). *Constructing organizational life: How social-symbolic work shapes selves, organizations, and institutions*. Oxford University Press.

Lounsbury, M., & Rao, H. (2004). Sources of durability and change in market classifications: A study of the reconstitution of product categories in the American mutual fund industry, 1944–1985. *Social Forces*, *82*, 969–999.

Marcus, G., & Luccioni, S. (2023, April 17). *Stop treating AI models like people*. Substack.com; https://garymarcus.substack.com/p/stop-treating-ai-models-like-people

Matz, S. C., Teeny, J. D., Vaid, S. S., Peters, H., Harari, G. M., & Cerf, M. (2024). The potential of generative AI for personalized persuasion at scale. *Scientific Reports*, *14*(1), 4692.

McKinsey & Co. (2024, May 30). *The state of AI in early 2024: Gen AI adoption spikes and starts to generate value*. McKinsey & Company. https://www.mckinsey.com/capabilities/quantumblack/our-insights/the-state-of-ai

Mesoudi, A., Whiten, A., & Laland, K. N. (2004). Perspective: Is human cultural evolution Darwinian? Evidence reviewed from the

perspective of *The Origin of Species*. *Evolution*, *58*(1), 1–11.
Mollick, E. (2024). *Co-intelligence*. Random House.
Navis, C., & Glynn, M. A. (2010). How new market categories emerge: Temporal dynamics of legitimacy, identity, and entrepreneurship in satellite radio, 1990–2005. *Administrative Science Quarterly*, *55*, 439–471.
Negro, G., Koçak, Ö., & Hsu, G. (2010). Research on categories in the sociology of organizations. In G. Hsu, G. Negro, & Ö. Koçak (Eds.), *Categories in markets: Origins and evolution* (pp. 3–35). Emerald Group Publishing Limited.
Orlikowski, W. J., & Scott, S. V. (2023). The digital undertow and institutional displacement: A sociomaterial approach. *Organization Theory*, *4*(2): 1–16. https://doi.org/10.1177/26317877231180898
Ozcan, P., & Santos, F. M. (2015). The market that never was: Turf wars and failed alliances in mobile payments. *Strategic Management Journal*, *36*, 1486–1512.
Pache, A. C., & Santos, F. (2013). Inside the hybrid organization: Selective coupling as a response to competing institutional logics. *Academy of Management Journal*, *56*, 972–1001.
Peterson, R. A., & Anand, N. (2004). The production of culture perspective. *Annual Review of Sociology*, *30*, 311–334.
Phillips, N., & Brown, J. L. (1993). Analyzing communication in and around organizations: A critical hermeneutic approach. *Academy of Management Journal*, *36*, 1547–1576.
Phillips, N., & Moser, C. (2024). The biological basis of the symbolic: Exploring the implications of the co-evolution of language, cognition and sociality for management studies. *Journal of Management Studies*. Advance online publication. https://doi.org/10.1111/joms.13037
Placani, A. (2024). Anthropomorphism in AI: Hype and fallacy. *AI and Ethics*. Advance online publication. https://doi.org/10.1007/s43681-024-00419-4
Porter, J. (2023, November 6). ChatGPT continues to be one of the fastest-growing services ever. *The Verge*. https://www.theverge.com/2023/11/6/23948386/chatgpt-active-user-count-openai-developer-conference
Power, M. (2022). Theorizing the economy of traces: From audit society to surveillance capitalism. *Organization Theory*, *3*(3), 26317877211052296.
Rao, H. (1998). Caveat emptor: The construction of nonprofit consumer watchdog organizations. *American Journal of Sociology*, *103*, 912–961.
Rao, H., Monin, P., & Durand, R. (2003). Institutional change in Toque Ville: Nouvelle cuisine as an identity movement in French gastronomy. *American Journal of Sociology*, *108*, 795–843.
Ruef, M., & Patterson, K. (2009). Credit and classification: The impact of industry boundaries in nineteenth-century America. *Administrative Science Quarterly*, *54*, 486–520.
Seaver, N. (2019). Captivating algorithms: Recommender systems as traps. *Journal of Material Culture*, *24*, 421–436.
Seeber, I., Bittner, E., Briggs, R. O., De Vreede, T., De Vreede, G. J., Elkins, A., Maier, R., Merz, A. B., Oeste-Reiß, S., Randrup, N., Schwabe, G., & Söllner, M. (2020). Machines as teammates: A research agenda on AI in team collaboration. *Information & Management*, *57*(2), 103174.
Shen, R., & Phillips, N. (2024). *Discourse and institutions: The role of digital technology*. University of California, Santa Barbara Working Paper.
Stark, D., & Broeck, P. V. (2024). Principles of algorithmic management. *Organization Theory*, *5*(2), 1–24. https://doi.org/10.1177/26317877241257213
Sun, Y., & Suthers, D. D. (2023). From affordances to cultural affordances: An analytic framework for tracing the dynamic interaction among technology, people and culture. *Cultures of Science*, *6*, 235–249.
Tommasel, A., & Menczer, F. (2022, September). *Do recommender systems make social media more susceptible to misinformation spreaders?* Proceedings of the 16th ACM Conference on Recommender Systems, Seattle, WA (pp. 550–555). https://doi.org/10.1145/3523227.3551473
Turing, A. M. (1950). Computing machinery and intelligence. *Mind*, *49*, 433–460.
Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., Kaiser, Ł., & Polosukhin, I. (2017). *Attention is all you need*. *arXiv preprint arXiv:1706.03762*.
Walsh, T. (2017). The singularity may never be near. *AI Magazine*, *38*(3), 58–62.
Walsh, T. (2024, January 19). AI paintings have been sold for huge sums. Can they make us feel the way art created by humans does? *Scroll.in*. https://scroll.in/article/1062081/ai-paintings-

have-been-sold-for-huge-sums-can-they-make-us-feel-the-way-art-created-by-humans-does

Walters, W. P., & Murcko, M. (2020). Assessing the impact of generative AI on medicinal chemistry. *Nature Biotechnology*, *38*, 143–145.

Warwick, K., & Shah, H. (2016). Passing the Turing test does not mean the end of humanity. *Cognitive Computation*, *8*, 409–419.

Whittle, A., Housley, W., Gilchrist, A., Mueller, F., & Lenney, P. (2015). Category predication work, discursive leadership and strategic sensemaking. *Human Relations*, *68*, 377–407.

Zhou, E., & Lee, D. (2023). Generative AI, human creativity, and art. *SSRN*. https://ssrn.com/abstract=4594824

Zuckerman, E. W. (1999). The categorical imperative: Securities analysts and the illegitimacy discount. *American Journal of Sociology*, *104*, 1398–1438.

## Author biographies

Nelson Phillips is the Christian A. Filipe Professor in the Department of Technology Management at the University of California, Santa Barbara. Professor Phillips' research interests cut across organization theory, innovation, and technology, and he has published widely for both academics and practitioners. He is the co-editor of *Innovation: Management and Organization* and series co-editor of the *Cambridge Elements in Organization Theory*. He was elected a fellow of the Academy of Management in 2021.

Sai Kalvapalle is an Assistant Professor of Strategy and Entrepreneurship at the Nova School of Business and Economics in Lisbon, Portugal. She earned her PhD from the Rotterdam School of Management in 2023. Her research interests are in organization theory and entrepreneurship, with a focus on communication across different organizational settings.

Mark T. Kennedy is an associate professor in the Business School at the Imperial College London. He received his Ph.D. from Northwestern University. His research focuses on the emergence of new markets and industries and the more basic building blocks of organizing, such as categories, identities, forms, and strategies.