Technovation 147 (2025) 103304

Contents lists available at ScienceDirect

Technovation

journal homepage: www.elsevier.com/locate/technovation

Algorithms in the room: AI, representation, and decisions about sustainable futures

Stuart Mills [a,b,*], Henrik Skaug Sætra [c]

[a] Department of Economics, Leeds University Business School, University of Leeds, UK
[b] London School of Economics, UK
[c] Department of Informatics, University of Oslo, Norway

ARTICLE INFO

Keywords:
Generative AI
Sustainable development goals
Representation
Decision-making

ABSTRACT

This article considers the role of generative AI technologies, such as large language models (LLMs), in promoting the views of underrepresented groups. We are specifically concerned with the role AI could play in encouraging powerful decision-makers—often leading politicians and businesspeople in Western nations—to consider the perspectives of underrepresented groups when making decisions about sustainable development.

Some suggest generative AI could offer decision-makers perspectives they had previously not considered, leading to more equitable and innovative policy approaches, and supporting several of the United Nations' Sustainable Development Goals (SDGs). We critique this perspective. Groups may be underrepresented in sustainable development decision-making because of individual cognitive and organisational information-processing limitations ('*omitted, but not opposed*'), and because of opposition which remains even if these limitations are overcome ('*opposed, whether omitted or not*'). We outline how these 'categories of omission' shape the opportunities and risks created by generative AI in representative sustainability.

1. Section 1: introduction

Representation is important in many domains – particularly when a diverse group of stakeholders have significantly different perspectives, values, and experiences. This article focuses on sustainable development as a case study for the use of generative artificial intelligence (AI) to help, or hinder, representativeness.

Sustainable development is a particularly relevant example because it tends to encompasse decisions involving economic, cultural, and social trade-offs. As such, effective sustainable development often requires thorough engagement with and representation of stakeholders to ensure that all relevant voices feel involved, and to enable decisions to be based on the broadest possible knowledge base. For instance, the European Commission's Corporate Sustainability Reporting Directive (CSRD) requires organisations to perform stakeholder analyses *and* include stakeholders in various decision processes. One example is in the performance of the mandatory double materiality assessment, which strongly encourages dialogue and consultation with relevant stakeholders (EFRAG, 2024).

Key areas of sustainable development often reveal wide-ranging and conflicting perspectives which must be resolved. Consider electric vehicles. Electric vehicle development unavoidably involves stakeholders ranging from indigenous communities in mineral rich areas such as South America and Australia; industrial and manufacturing centres in areas such as East Asia; and consumer markets, typically those of North America, Europe, and increasingly, China (Marx, 2022). Decisions about this industry, and others, therefore, involve judgements which mix economic, social, cultural (and more) dimensions. In doing so, conflict around who makes decisions, and what information is used in decision-making, can arise (Sætra, Mills and Selinger, 2025), and these conflicts must often be resolved for development progress to be made.

However, as some scholars have documented, in many instances the 'solution' to these conflicts of many stakeholders in sustainability and development is to ignore marginalised groups (Ostrom, 1990, 1996). For instance, indigenous groups have long protested the absence of their voices from global climate discussions, such as the various Conference of the Parties (COP) gatherings; or, if included, their diminished and often 'token' role (Lakhani, 2021). Beyond indigenous representation, female participation in delegations for COP27 was only around 34 % (Stallard, 2022), despite women and girls predicted to suffer more from climate

This article is part of a special issue entitled: AI x SDGs published in Technovation.

* Corresponding author. University of LeedsMaurice Keyworth Building, Leeds University Business School, Leeds, LS2 9JT, UK.
E-mail address: s.mills1@leeds.ac.uk (S. Mills).

https://doi.org/10.1016/j.technovation.2025.103304
Received 29 September 2024; Received in revised form 23 June 2025; Accepted 29 June 2025
Available online 3 July 2025

change than men and boys (Dunne, 2020).

Underrepresented groups often suffer worse outcomes arising from the 'negative externalities' of ignored perspectives (Stiglitz, 2024). For instance, the so-called 'medical research gender gap' describes the consequences of women, and women's health, being consistently underrepresented within medical research (Merone et al., 2022), leading to worse health outcomes for women to this day (Maas and Appelman, 2010; Neville, 2024). Such outcomes run contrary to various stated aims within the United Nations' (2015) sustainable development goals (e.g., SDG 5, 9, 10, and 16). A lack of representation can also lead to *better* solutions being missed—a key argument found within discussions of diversity, equity, and inclusion (Sætra, 2024a). Underrepresented groups often contribute novel ideas which *improve* the quality of decision-making and foster more innovative approaches (Ostrom, 1990). For instance, the Intergovernmental Panel on Climate Change (IPCC) has recently acknowledged the role of indigenous knowledge to *enhance* the methodologies used for understanding and measuring the effects of climate change (Mohamed et al., 2022).

How can underrepresented groups be better represented in decision-making? One solution may be generative AI. These systems are generative insofar as they produce outputs which can be informationally greater than the inputs (prompts) given to them (Sætra, 2023). Some have suggested generative AI could generate views and perspectives that decision-makers might otherwise miss, expanding the information available to decision-makers and leading to better decisions (Agnew et al., 2024). Others suggest generative AI can combine existing information in novel ways, encouraging decision-makers to see problems (and solutions) differently (Bouscherry et al., 2023), leading to better outcomes in terms of reduced conflict, and fewer negative externalities. This article critically examines such propositions.

Drawing on the sustainable development as a backdrop, we propose two broad categories for why different perspectives come to be underrepresented in decision-making, which we call 'categories of omission.' The first we call 'omitted, but not opposed,' where perspectives are omitted due to the cognitive limits of decision-makers and decision-making processes. However, these perspectives would not be omitted if these limits were overcome. The second we call 'opposed, whether omitted or not.' Here, while cognitive limits will still exist, omission persists because decision-makers oppose inclusion for some (material or ideological) reason. These categories reflect the individual and organisational 'costs' of incorporating information into decision-making. They also capture the important distinction between *providing* information to decision-makers, *versus* changing *who* is deciding. We draw on various perspectives from behavioural science (Simon, 1955, 1956, 1981; Tversky and Kahneman, 1974), organisational studies (Cohen et al., 1972; Cyert and March 1992; March and Simon, 1993; Simon, 1997), and political agenda setting (Hoefer, 2022; Kingdon, 2010; Lindblom, 1959) to inform our discussion.

With this framework, we evaluate the opportunities and challenges generative AI creates in relation to representation and sustainable development. We argue that the conceptual and practical considerations of using generative AI for sustainable development decision-making depends on a deeper understanding of why groups are underrepresented in the first instance.

The structure of this article is as follows. Section 2 develops the foundation of our categories of omission. Section 3 reviews two literatures which are emerging around the question of AI and representation. These are the algorithmic fidelity and AI innovation literatures. Section 4 uses our categories of omission to critique the arguments found in these literatures. We divide our critique into four sections, focusing on the opportunities and risks generative AI creates, given each category of omission. Section 5 offers some recommendations and perspectives on future research, before Section 6 concludes.

2. Section 2: why are perspectives omitted?

To appreciate how generative AI might promote more inclusive processes related to sustainability, and the problems which may also arise, it is important to understand why different perspectives can be omitted by decision-makers. This is a broad question, and the framework we develop draws on various insights into individual decision-making, organisational behaviour, and political agenda-setting. It is thus applicable beyond the domain of sustainable development. This framework *is* appropriate to the domain of sustainable development insofar as decisions occur at various organisational levels and in conjunction with different groups and environments. For instance, the COP gatherings typically seek to arrive at an international agreement on carbon emissions; each nation then has flexibility in how they meet this agreement (including whether they do *not* meet the agreement); the national agenda is influenced by that nation's key decision-makers (e.g., politicians); whose agenda may then be implemented by institutions consisting of civil servants who, at an individual level, influence 'on-the-ground' or 'street-level' outcomes (e.g., Herd and Moynihan, 2018).

The framework we develop seeks to recognise this broad decision-making infrastructure while remaining parsimonious for the purposes of usability. We propose two categories of omission: 'omitted, but not opposed,' and 'opposed, whether omitted or not.'

These categories follow from observations about decision-making given, independently, by Simon (1997) in his work on organisational behaviour, and Kingdon (2010) in his work on agenda-setting in policy. Simon (1997) emphasises how an organisation's *value premises* (e.g., what *ought* to be done?) establishes constraints on individuals who make choices according to *factual premises* (e.g., what *can* be done, given what *ought* to be done?). Kingdon (2010) outlines how elected officials set the *governmental agenda* (e.g., the politician's commitments) which, in turn, determines the *decision agenda* (e.g., how can these commitments be achieved?). These distinctions are helpful for exploring why perspectives might be omitted from decision-making. Omission around *factual premises* or the *decision agenda* may typically be understood in terms of cognitive limitations, as decision-makers must evaluate different ways of achieving some broad objective. Omission around *value premises* or the *governmental agenda*, however, will more often reflect the personal interests, beliefs, and capabilities of key decision-makers.

Our two categories of omission loosely align with these two 'levels' of decision-making, though imperfectly. For instance, organisational capacity (factual premises) may constrain what the organisation *can do*, thus shaping what the organisation practically *ought to do* (Simon, 1997). We elaborate on these dynamics, below. Nevertheless, while imperfect, our framework offers a structure for evaluating use-cases emergent in the literature (Section 3) in contrast with broad decision-making dynamics, ultimately to explore reasonable avenues of critique (Section 4) and develop worthwhile recommendations (Section 5).

2.1. Omitted, but not opposed

Cognitive limitations may lead some perspectives to be overlooked by decision-makers. People are boundedly rational (Simon, 1997), with limited information processing capabilities (Simon, 1955) which are often influenced by one's environment (Simon, 1956). Highly complex decisions, for instance those that must balance global stakeholder perspectives, will often be difficult for individuals to navigate, owing to the large amount of information involved, and that information's complexity (Simon, 1997). In such instances, people often simplify decisions, sometimes through effective search heuristics (e.g., Simon, 1981, 1955), and sometimes through heuristics which lead to biases and other decisional errors (e.g., Tversky and Kahneman, 1974; for a recent review focusing on climate and sustainability, see Wouter Botzen et al., 2025).

These cognitive processes often enable *some* decision to be made, but

usually at the expense of an alternative perspective. An important example is the availability heuristic, a strategy for simplifying complex information processing tasks by using information which is easily brought to mind (Tversky and Kahneman, 1973). For instance, a decision-maker from a developed country may believe investments in electric vehicle infrastructure are most important for sustainability because they can easily imagine a developed city laden with combustion engine vehicles. The immediate availability of a sustainability problem (combustion vehicles) and solution (electric vehicles) will likely help the decision-maker navigate their potential choices. Yet, the trade-off of this heuristic is that other elements of this investment decision may be overlooked. For instance, investment in electric vehicle infrastructure may intensify the environmentally damaging extraction of minerals needed for said vehicles (Marx, 2022). This trade-off reflects the role of cognitive limitations at the level of the decision agenda or in terms of factual premises (e.g., how can sustainable investment be undertaken?). The decision that investment *ought to be* undertaken is a separate decision arising at the level of the governmental agenda and reflecting value premises.

Further complicating matters is the environment in which boundedly rational decision-makers inhabit (Simon, 1956). For instance, the simple act of viewing a decision as an *investment* in sustainable development may drive attention towards one set of facts, and away from an alternative—but still relevant—set (Kingdon, 2010; Simon, 1997). An economist may view a sustainable development decision as an investment, with payoffs (economic and political). But a climate scientist may view the same decision in terms of carbon emissions and environmental impact. An indigenous population may view the decision as a reparation for historical damages, or an appropriation of their resources (de Beukelaer, 2023). In recent years, many climate scientists have been reflecting on this problem as they come to recognise that their scientific expertise often does not accord with the language, outlook, and perspective of politicians, impeding the effectiveness of scientific advocacy (Howarth et al., 2020).

When perspectives are omitted because of cognitive costs, there may be little actual *opposition* to the omitted perspective. Instead, omission is based on a lack of consideration owing to one's role and personal bias on a topic (Cyert and March 1992; Kingdon, 2010; Moore et al., 2024; Simon, 1997; Wouter Botzen et al., 2025). An economist born and raised in a wealthy nation is unlikely to view many sustainable development decisions through the eyes of an indigenous conservationist raised in a developing nation, and *vice versa*. Often, different groups will approach the same problem or decision with different sets of facts and values determined by their backgrounds, specialist knowledge, experience (etc., Sen, 2002; Simon, 1997), and no 'objective' way of evaluating or dismissing these perspectives will be available (Cyert and March 1992; Hoefer, 2022; Polanyi, 2005). For instance, one recent review of politicians' engagement with climate scientists found that those who expressed an intrinsic interest in sustainability issues tended to have personal experiences of climate-related events prior to their political careers (Moore et al., 2024). The availability of these experiences likely encourages these politicians to engage with, rather than ignore, climate scientists. Similar observations have been raised in investigations of race and whiteness as an organisational default (Sue, 2006).

In such instances, the decision-maker will often focus on those perspectives which present a personally intuitive narrative, without necessarily passing judgement on the substance of alternative, though overlooked, perspectives (Kingdon, 2010). As Kingdon (2010) notes, many perspectives are overlooked not because of explicit opposition, but rather from this chaotic confluence of factors which at one moment may favour one perspective, and at another moment, another (also see Cohen et al., 1972; Lindblom, 1959; March and Simon, 1993), as decision-makers seek to balance and ration their limited cognitive resources (Simon, 1997).

Thus, some perspectives are omitted by decision-makers not due to any specific opposition to those perspectives, but simply because of the cognitive costs which arise from considering more perspectives, and the simplifying heuristics often employed to manage this informational burden. These omissions constitute a problem of underrepresentation we refer to as *omitted, but not opposed*. As such, if one could overcome these cognitive limits, omitted perspectives may very well be integrated into a more representative decision-making process.

2.2. Opposed, whether omitted or not

By contrast, decision-makers will at times simply oppose certain perspectives. This may be best understood in terms of the value premises which reflect what key decision-makers believe, desire, or think *ought to* happen.

For Kingdon (2010), the role of key political actors cannot be ignored when considering opposition, as these individuals (e.g., elected officials) to determine what issues *ought to be* considered, what meets the criteria of evidence, and so on. These decisions, in turn, set boundaries on which perspectives will be considered, and which will be overlooked (also see Simon, 1997). For instance, reparations for indigenous groups may entail politically unpopular financial compensation, leading political actors to omit perspectives which advocate for reparations by setting a different policy direction. Political factors compound further when decisions involve conflicts over value systems (Ostrom, 1990), as establishing a value system is a key decision taken by authoritative decision-makers (Simon, 1997). For instance, financially compensating indigenous groups assumes these groups subscribe to an instrumental view of value (e.g., that land has a price) rather than an intrinsic view (e.g., that nature is itself is valuable). If a decision-maker can (or will) only entertain one value-system, those whose views are based on a different system are likely to be ignored or diminished.

As above, Moore et al. (2024) report that a key driver of political engagement with climate change policy is the politician's intrinsic motivations, given their personal experiences. They also note, though, that climate rises up the political agenda for those politicians whose constituents express particular interest and concern in such matters. In either instance, Moore et al. (2024) suggest the decision to focus on, or sideline, climate-related policies is directly tied to the personal motivations of elected officials.

Yet, while key decision-makers set agendas to *achieve* some outcome (e.g., re-election), perspectives may also be ignored in an effort to *avoid* some outcome, typically a conflict. Grove (1997), writing from a management perspective, argues perspectives may be ignored because key decision-makers do not want to engage with tough issues, such as moving out of markets or laying off employees. Simon (1997), too, discusses this aspect of organisational decision-making, emphasising the role of dilemmas and the frequent desires of people to avoid undesirable choices. Specifically, Simon (1997) argues people will often avoid information which demands a confrontation with a past, bad decision, as well as information which might commit one to negative future outcomes. These descriptions thus represent a kind of 'ostrich' strategy of burying one's head in the sand. For Grove (1997) and Simon (1997), where perspectives create conflict, these perspectives may be ignored because one's objective is to *avoid* conflict.

Though, whether to achieve some objective, or to avoid some outcome, it would be a mistake to contend that those who set the governmental agenda or establish the value premises of an organisation have supreme authority over what is considered, and what is ignored. Kingdon (2010) and Simon (1997) both emphasise the role of what might be dubbed organisational costs and coordination costs.

A key decision-maker, such as an elected official, may have a strong role to play in determining which outcomes are to be pursued, and thus which perspectives are to be considered. But the outcomes which are pursuable are also constrained by the resources available to organisations, and the individual cognitive resources of people. For instance, both Kingdon and Simon discuss the role of budgets and past decisions in determining future courses of action. Perceived or actual economic

constraints may inspire opposition to some perspectives simply because of the costs associated with them. Such costs are often tied to past decisions. For instance, perspectives which advocate mass pedestrianisation of cities are frequently opposed because previous investments in motor vehicle transport entail substantial costs to reverse (Marx, 2022; Shreedhar et al., 2024). Economic costs, typically discussed as sunk costs, are often cited as reasons for not divesting from fossil fuel infrastructure (Pettifor, 2019; Stiglitz, 2024).

Crucially, none of these accounts of organisational behaviour point to a *lack of information* as the reason for omission. All emphasise how the range of perspectives which can be considered is frequently constrained by the agenda-setting decisions of key decision-makers, in conjunction with the capabilities and resources of organisations themselves. In such instances, providing people with more information will not necessarily result in representative outcomes. Writing on international trade deals and economic cooperation, Stiglitz (2024, p. 242) has described this phenomenon as the "façade of inclusiveness." He notes that, "Developing countries have demanded to take part in crucial global agreements because they have realised that if you don't have a seat at the table, you may be on the menu. But having a set at the table isn't enough. Too often, their microphone has been effectively turned off, and no one is listening."

Thus, perspectives may be omitted because they conflict with agendas, interests, and resource constraints, or because of a desire to avoid conflicts and difficult decisions, or both. These reasons do not forestall the possibility that decision-makers are also subject to various cognitive and organisational constraints which lead to oversights, as described above. Yet, what distinguishes those *omitted, but not opposed* omissions from those discussed presently is that even if individual cognitive limitations were to be overcome, decision-makers would *still* not consider some perspectives due to other factors, such as personal interests and organisational resource limitations. Thus, these omissions—those which are *opposed, whether omitted or not*—represent a distinctly different category of omission to *omitted, but not opposed* omissions.

3. Section 3: algorithmic fidelity and AI innovation

How might generative AI ameliorate problems of representation? Two emerging literatures in this space are the 'algorithmic fidelity' and 'AI innovation' literatures. The former examines how generative AI can simulate populations to inform actual decision-making (Agnew et al., 2024). The latter examines how AI can inspire innovations by recombining ideas in novel ways (Bouschery et al., 2023; also see Mariani et al., 2023). As such, both literatures suggest AI can support more holistic decision-making through giving new information and perspectives to decision-makers. It is from this perspective that these literatures contribute to a discussion of generative AI within representative decision-making.

3.1. Algorithmic fidelity

The term 'algorithmic fidelity' comes from Argyle et al. (2023). They (p. 339) define algorithmic fidelity as, "the degree to which the complex patterns of relationships between ideas, attitudes, and sociocultural contexts within a [large language] model accurately mirror those within a range of human subpopulations." Algorithmic fidelity arises because, "these language models do not contain just one bias, but *many*" (p. 338, original emphasis), allowing for simulations of many different people in ways which mirror comparable populations—what is known as *silicon sampling*. For Argyle et al. (2023, p. 338, original emphasis), algorithmic fidelity arises because generative AI models are often "biased both toward *and* against specific groups and perspectives in ways that strongly correspond with human response patterns along fine-grained demographic axes."

Argyle et al. (2023) outline some criteria for algorithmic fidelity. Firstly, they note that simulating single individuals may often lead to responses which are incoherent when compared with a demographically similar human. Secondly, they emphasise the importance of interpreting outputs in *context*. This is to say, outputs should be consistent with inputs in terms of tone and content. A response which may appear reasonable in isolation is unlikely to be an accurate simulation if there is no obvious coherence between it, and the input context.

Several studies have investigated algorithmic fidelity and silicon sampling. Consumer behaviour scholars, for instance, have demonstrated that silicon samples created using GPT-3.5 show comparable product preferences and willingness-to-pay metrics as consumer groups (Brand et al., 2024). These samples also produce comparable accounts of experiences and opinions about different products (Hämäläinen et al., 2023). In terms of economic games, silicon samples created through GPT-4 have also been shown to be consistent with human behaviour (Aher et al., 2023; Mei et al., 2024).

Others report mixed results. Lee et al. (2024) show that GPT-4 can generate synthetic populations which accurately simulate presidential voting behaviours. However, further contextual priming is needed to accurately simulate policy opinions about global warming and related environmental policies. Furthermore, Lee et al. (2024) find that the silicon sample fails to accurately simulate the concerns of Black Americans about global warming. In a similar study, examining a broader set of political opinions, Hwang et al. (2023) find comparable results. Santurkar et al. (2023) find poor alignment between large language model (LLM) simulation and political opinion polling across 60 different demographic groups in the US. Furthermore, they argue current LLMs are likely not trained on sufficient data to accurately simulate some demographic groups, such as older individuals. Similarly, Shrestha et al. (2025) report poor alignment between GPT-4 simulations and human responses when examining policy opinions, particularly in silicon samples of non-WEIRD (western, education, industrialised, rich, and democratic) people. This is particularly relevant to many sustainability-related issues, where non-WEIRD peoples are often the minorities and stakeholders whose interests are affected by the decisions of both political and corporate decisions-makers.

A lack of representative training data within current LLMs might also be shown in an interesting study by Gmyrek et al. (2024). They find that GPT-4 evaluations of various occupations (in terms of prestige, social value, etc.) are comparable to human evaluations only when high-level categories are used (e.g., doctor). When low-level categories are used (e.g., oncologist), the simulation accuracy falls. This is because there are fewer examples of the low-level categories, causing the more numerous high-level examples to dominate the final output (also see Sorensen et al., 2024). Peterson (2025) argues that LLMs may *always* struggle to capture minority views due to their engineering. As LLMs are designed to be aggregators of training data, Peterson (2025) suggests the long-tails in these data—anomalous or infrequent datapoints—are inevitably trained *out* of the model, while the average of the dataset becomes overrepresented.

Finally, Amirova et al. (2024) investigate silicon sampling in terms of qualitative data *and* (quantitative) survey data. They use GPT-3.5 to simulate interviews, which are then compared to interviews of people. Amirova et al. (2024, p. 1) find LLM simulations to be "strikingly similar" to human responses in terms of broad interview themes. However, using qualitative methods, they determine that the simulated interviews differ substantially in terms of interview structure, tone, and some elements of language style.

3.2. AI innovation

Discussion of AI as a tool for innovation has a decades-long history. In one early discussion of the possibility of creative AI, Boden (1998) argued that creativity broadly consists of achieving novelty through recombination of familiar ideas to realise previously unimagined approaches. They suggested these processes could feasibility be

undertaken by an AI system. Simon (1997, 1981), too, presented similar arguments—though more often from an information management perspective.

More recently, Bouschery et al. (2023) have iterated on the AI innovation discussion by discussing the potential role of generative AI. They draw on the 'double diamond' model of innovative problem solving (Howard et al., 2008; see Fig. 1) to show how generative AI can be incorporated into the innovation process. This model conceptualises innovation as a process occurring over two 'spaces' and four stages. These spaces are the 'problem space' where (1) problems are articulated and (2) important problems selected, and the 'solution space' where (3) solutions are generated and (4) important solutions selected. Bouschery et al. (2023) argue that generative AI can expand the range of ideas considered at the (1) problem articulation and (3) solution generation stages, as do Brem et al. (2023).

Si et al. (2024) provide some evidence to support this argument. In an experiment where researchers were asked to blindly evaluate research hypotheses generated by humans and generative AI systems, the latter were found to be more novel than the former, though less practically feasible. Kakatkar et al. (2020) also draw on the double diamond model. Like Bouschery et al. (2023), they emphasise how AI can enhance stages (1) and (3). However, they also argue AI can support stages (2) and (4). Using case study evidence of AI being used in innovation management, Kakatkar et al. (2020) show AI often serves as a) a descriptive tool; b) a diagnostic tool; c) a predictive tool; and d) a prescriptive tool. While these functions can support articulation of problems and solutions; AI prediction and diagnostics could support *selection* stages. They thus emphasise the role of AI as an information management tool within the double-diamond model, as well as a generative tool, allowing not only more problems and solutions to be generated, but for problems to be better defined, and better-suited solutions selected.

This connects to another emerging theme in the AI innovation literature, which is the use of generative AI for forecasting and predictive exploration (e.g., Schoenegger et al., 2024). Füller et al. (2022) survey 150 innovation managers currently using AI, and find managers predict AI will be most useful in predicting and exploring changing consumer trends. Bilgram and Laarmann (2023) focus specifically on LLMs. They argue that generative AI will accelerate innovation practices, providing illustrative evidence of GPT-3 supporting a PESTEL analysis of the automotive market; an analysis of consumer needs in this same market; a simulation of consumer profiles; and simulation of consumer feedback. These arguments align with other arguments that generative AI can support innovation through expanding problem and solution spaces via novel combinations (e.g., Boden, 1998; Bouschery et al., 2023; Brem et al., 2023; Haefner et al., 2021; Kakatkar et al., 2020; Si et al., 2024).

3.3. Limitations

Both literatures point to opportunities and limitations of AI technology.

The algorithmic fidelity literature shows promising predictive results in some domains. There are also several examples of misalignment between silicon samples and human responses, which reflect a multitude of nuances. For instance, the literature—being nascent—shows methodological inconsistencies in terms of evaluating accuracy to human samples. Equally, many studies available at the time of writing draw on older LLMs and might not reflect samples created using more advanced generative AI systems. For the purposes of this discussion, this literature shows the opportunities for generative AI to supplement decision-maker knowledge of underrepresented groups, but that generative AI models may also lack adequate representation of those groups. This might still lead to a situation in which the perspectives and ideas are expanded by using AI, but it could be an expansion largely in line with majority or privileged positions. This could lead to new discoveries and innovations but would not resolve issues related to exclusion or underrepresentation.

The AI innovation literature, building from the double-diamond model, offers a practically useful conceptual understanding of how generative AI can support innovative decision-making. While the novelty of generative AI ideas may be encouraging insofar as one wishes to foster innovation and draw attention to new (perhaps overlooked) perspectives, the finding that many generative AI ideas lack feasibility (compared to human ideas) may point to potential challenges around, say, budgetary and resource constraints, which limits the practical usage of these technologies. As with the algorithmic fidelity literature, though, such limitations are subject to current technological capabilities and

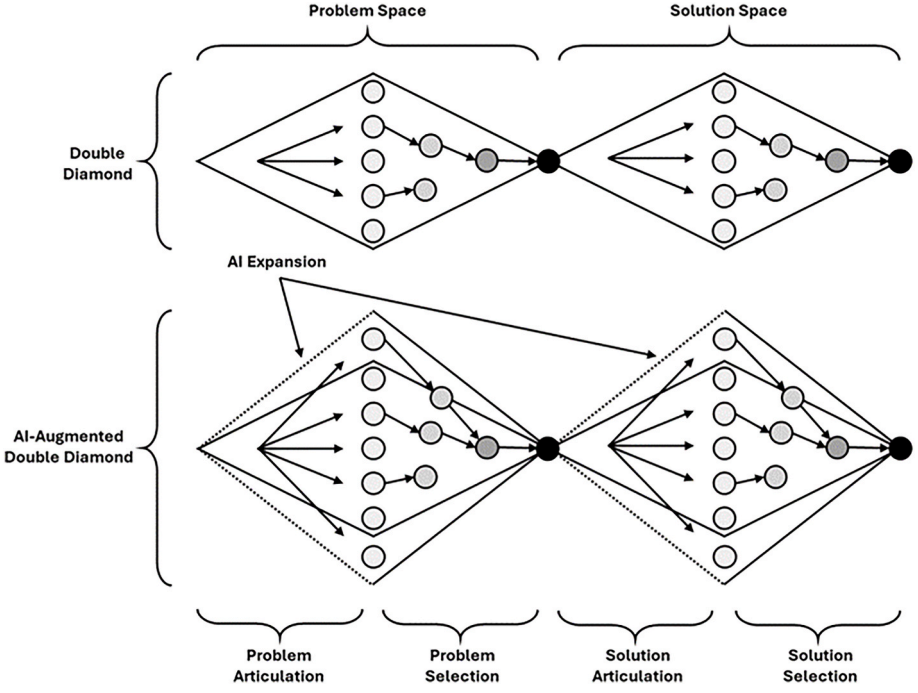

Fig. 1. The 'Double Diamond' Model incorporating the use of AI. Figure adapted from Bouschery et al. (2023).

reflect the developing best-practice methodologies of an emerging field of study.

It would be unwise to conclude that current limitations found in either literature define the limits of generative AI indefinitely. Equally, applications generative AI today are necessarily applications of *today's* technology, and thus any analyses or recommendations must build from current capabilities. Our review of these two literatures point to important areas of critique around the use of generative AI for more representative decision-making, to which we now turn.

4. Section 4: generative AI for representative sustainability?

The algorithmic fidelity and AI innovation literatures establish arguments for how generative AI can support greater representation in various domains. We now contrast these arguments against our categorises of omission. Table 1 summarises the arguments developed in this section, which we set against a discussion of the implications for the SDGs. Table 1 is unlikely to be exhaustive, given both the evolving nature of AI technologies and the complexity of representation challenges. We comment on an important limitation—how technology and social problems change *over time*—at the end of this section.

4.1. Opportunities when 'Omitted, but not opposed'

Opportunities likely exist to promote more representation in sustainability related decision-making when omitted perspectives are not in principle opposed by decision-makers. As above, people suffer from cognitive limitations which mean we are unable to synthesise all relevant information, while being pre-disposed to information which aligns with our familiarities and backgrounds (Simon, 1955, 1956, 1997; Tversky and Kahneman, 1974). The proposed capabilities for generative AI to simulate the perspectives of groups whom decision-makers may otherwise overlook is potentially a powerful innovation in this regard, as is the potential for generative AI to develop novel ideas to interrogate the pre-existing policy positions of decision-makers (Cantens, 2025).

An obvious retort is the following: if a decision-maker is sufficiently knowledgeable of their ignorance to know generative AI would help them make more representative decisions, why not simply *invite* underrepresented groups to participate in the decision-making process? However, the feasibility of this suggestion depends on the context of the decision being made. In some instances, such as elections, democracies do (in principle) indeed invite everyone to participate. Here, generative AI is likely a poor—probably *unacceptable*—replacement for participation by all (willing) members.

Yet, in other scenarios, various groups may not be able to participate for several reasons. For instance, Specian (2023) argues AI systems could provide aid to decision-makers ('machine advisors'; Specian, 2024) when relevant policy experts are unavailable, due to time and economic cost constraints. Time and cost constraints are also reflected in commercial interests in algorithmic fidelity and AI innovation approaches (Bouschery et al., 2023; Brand et al., 2024; Hämäläinen et al., 2023). A recent review of the algorithmic fidelity literature identified cost-savings to be the most commonly cited benefit of AI simulation (Agnew et al., 2024).

From a sustainability perspective, the substantial uncertainty surrounding various decisions (such as climate policy), coupled with the large scales which are often involved, may make *actual* experimentation with multiple policy approaches infeasible. Therefore, using AI to simulate and predict different policy outcomes may have practical benefits, and support the SDGs. Furthermore, uncertainty around sustainability may demand an immediate policy response, which generative AI might be able to support, while for poorer stakeholders (e.g., poorer countries), or for stakeholders lacking capabilities (e.g., surveying infrastructure), generative AI might be an important tool for overcoming organisational constraints on effective sustainability actions (Sætra, 2022).

Generative AI could enable many different groups to participate in decision-making, in a manner of speaking (Specian, 2023). For instance, LLMs can function as efficient tools for summarising many perspectives, and handling a volume of information that people would struggle to attend to. Such information would likely be overlooked by people without technological support (Tessler et al., 2024; also see Simon, 1987a, 1987b). Generative AI may allow more perspectives to be integrated into a smaller pool of information which key decision-makers actually use, supporting greater representation and allowing novel insights to be retained. For instance, Aonghusa and Michie (2020) report on experimental trials of an AI system for summarising public health research materials. This system makes policy predictions based on these materials, which are given to policymakers to support their policy choices and implementation strategies. Such uses may have positive implications for innovation (SDG 9) and might further strengthen institutions by enabling them to function more efficiently, making better use of existing knowledge and expertise (SDG 16).

In summary, when representation challenges are '*omitted, but not opposed*,' generative AI may help by simulating perspectives which would otherwise be overlooked, and summarising information so as to reduce the likelihood of it being overlooked. In both instances, generative AI could thus be used to ameliorate the cognitive and organisational limits on information processing—the *causes* of omission within this category.

Table 1
Opportunities and risks of generative AI, given categories of omission.

	Categories of Omission	
	Omitted, but not opposed	Opposed, whether omitted or not.
Opportunities	Generative AI can generate novel ideas and offer information about previously ignored perspectives, overcoming various cognitive biases and enhancing the quality of information available to decision-makers. Furthermore, it could help decision-makers synthesise large amounts of information into a decision by creating practically useful summaries of large surveys of public opinion and research, allowing more voices to be incorporated into the decision-making process.	Generative AI could help underrepresented groups to produce evidence/information/materials which promote their perspective *without* relying on the consent of established decision-makers. In this sense, generative AI could be a tool for 'democratising' decision-making processes by empowering groups who at present lack the resources to challenge established power structures.
Risks	Generative AI must generate outputs which are accurate and meaningfully reflect relevant stakeholders. Outputs must also be feasible given current constraints. Current technological challenges, such as AI hallucinations and the tendency towards the average, may mean these systems fail to capture/simulate the true plurality of a population (though technical developments may overcome some of these problems). Decision-maker ignorance of the system's accuracy and the feasibility of outputs make verifying the fidelity of an AI system difficult.	Generative AI tackles omission by ignorance much more than omission resulting from structural opposition or constraint. Some perspectives may be omitted *regardless* of their immediate availability. In these instances, generative AI offers little recourse to representation problems. Indeed, it may even exacerbate existing inequalities in representation by allowing decision-makers to *appear* to be tackling representation problems, when they are not.

4.2. Risks when 'Omitted, but not opposed'

Yet, as Specian (2023) notes, technologies which offer apparent benefits rarely do so without various, often underappreciated, risks. An important risk for the present discussion is the risk of inaccuracies in generative AI systems. The 'accuracy' of an AI simulation depends upon the broadness of the simulation and the methods used to evaluate it (e.g., Amirova et al., 2024; Gmyrek et al., 2024). Generative AI may be useful for enabling a broader *consideration* of underrepresented perspectives, but less useful when seeking more detail and nuance around specificities, owing to the relative scarcity of training data about these perspectives compared to 'higher level' knowledge areas (Peterson, 2025). In areas like sustainable development, where significant uncertainties (around, say, climate) are common, the risk of inaccuracies is also pertinent (Sætra, 2022). Yet, spotting inaccuracies reveals a fundamental challenge: if one is so ignorant about a group or perspective that one needs a tool such as generative AI, one is likely ill-equipped to spot AI inaccuracies and correct them. This means representation problems are prone to remain while the use of AI can change the character of these problems.

If one cannot spot AI inaccuracies, one may develop false confidence in the AI system. An antecedent to this argument is the critique of 'experts' and 'expertise' in democratic decision-making (Coeckelbergh, 2025; Feyerabend, 1978; Landemore, 2022; Specian, 2023). The notion of *epistemic dependence* occurs when people lack requisite knowledge about a decision, leading them to depend upon those who do have such knowledge (Coeckelbergh, 2025; Hardwig, 1985). For instance, politicians setting sustainability policies will often depend upon the insights of scientists when evaluating potential courses of action. But critics of epistemic dependence in democratic societies argue (amongst other points) that 'experts' frequently lack knowledge which may actually be relevant to the decision being debated (Ostrom, 1990). For instance, a scientist may know much about developing a renewable energy project, but lack the local knowledge to accurately determine how this project will impact a particular local community. Nevertheless, owing to the scientist's authority as an 'expert,' their testimony may be relied upon more than the testimonies of those with local (and potentially more valuable) knowledge.

A similar problem may arise from generative AI, in several ways. Firstly, confidence in the 'expertise' of generative AI may cause one to not question an AI generated output—what is known as *automation bias* (e.g., Alon-Barkat and Busuioc, 2023). For instance, if an AI output misses a relevant stakeholder or idea, but so too do decision-makers, the latter may believe they have considered all relevant details when in fact they have not. Secondly, decision-makers may fail to spot *mis*-representations of previously overlooked perspectives. If these misrepresentations are not identified, one may again become convinced that representation problems have been resolved, while decisions become *less* representative as they are made to support a group or perspective which does not exist.

Beyond accuracy, there is also the important question of what representation should mean. Generative AI may be able to supply decision-makers with information about novel perspectives. But for some, representation will not simply mean *consideration*, but *participation* (Sætra, 2020, 2024b). Thus, sidelining concerns about accuracy, as accepting that generative AI may support decision-makers in thinking more representatively does not by itself mean that decision-making *processes* will be considered representative or inclusive by groups who might still find themselves excluded.

4.3. Opportunities when 'Opposed, whether omitted or not'

Interestingly, it is from this epistemic dependence critique that Specian (2023) develops a compelling defence of generative AI systems in representative decision-making. Specian (2023) argues that generative AI systems are often much more amenable to everyday people than expert testimony is. AI outputs can simplify and de-jargon language and complex ideas. This can make these ideas more accessible to people, including underrepresented groups who—owing to their underrepresentation—may suffer other discriminations, including in access to education. Furthermore, generative AI systems are readily accessible to people with a computer and an internet connection (at the time of writing). They also lack the institutional reputations of experts, which often intimidate non-experts. As such, Specian (2023) speculates that ordinary people will be more likely to probe, interrogate, and debate AI outputs. Such interactions may even enable people to *feel* more involved in deliberative processes.

More specifically, though, when omission is because of decision-maker opposition, these arguments highlight a potential opportunity to overcome this omission and challenge institutional power (relating, in turn, to SDG 16). For instance, acquiring expertise can be expensive, as experts have limited time and highly valued skills. Organisations such as Exxon Mobil have the economic resources to hire experts to produce various materials to influence decision-makers (as well as directly lobby policymakers; Oreskes and Conway, 2010), while rival groups, such as indigenous communities, are unlikely to have the same level of access to experts, or resources to do so. Exxon Mobil, potentially having interests opposed to those of indigenous groups, is unlikely to advocate for the interests of these groups. As such, policymakers are likely to omit indigenous perspectives too, either because of the immediate availability of rival perspectives (omitted, but not opposed) or because of their aligned political interests with rival groups (opposed, whether omitted or not; Kingdon, 2010).

While generative AI cannot overcome *all* of the forces at play in this hypothetical (e.g., lobbying), the implications of which we consider in the following subsection, it *may* provide indigenous groups (in this instance) access to expertise they would otherwise struggle to acquire. In some instances, generative AI may even support the *communicating* of marginalised views to entrenched decision-makers by repackaging marginalised views in an amenable language and style. Generative AI may be a compelling mediator of discourse between experts and non-experts, or governors and the governed, and this could give underrepresented groups means to challenge their exclusion from decision-making processes (Specian, 2023). The potential distribution of accessible expertise brought by generative AI, coupled with the technology's potential to bolster the challenge underrepresented groups can make to their underrepresentation, can both feasibly contribute to SDG 16.

In summary, when perspectives are omitted because of opposition from decision-makers, either directly (e.g., developed nations in global agreements) or indirectly (e.g., through lobbying by opposing groups), using generative AI to provide decision-makers with the omitted perspective is unlikely to resolve the representation challenge underrepresented groups face. Ultimately, the *cause* of omission is not a *lack* of information, but structural impediments. Generative AI could *potentially* support better representation in decision-making, and further the SDGs, insofar as it can be used as an accessible tool for challenging these very impediments. By making stakeholder input (even if synthetic) so easily available, it is likely to become a natural part of best practice.

4.4. Risks when 'Opposed, whether omitted or not'

However, a major critique of this argument is that similar claims have been made about new technologies in the past, and such claims have rarely been proven right. For instance, Morozov (2011) has argued that various claims about the internet 'democratising' access to information, and thus empowering oppressed peoples to challenge their oppression, have not come to pass. Instead, Morozov (2011) asserts that the internet has often been co-opted by oppressing groups to maintain and entrench their power over others. The historian and social critic Ivan Illich (1981)) has presented a similar argument regarding what was perhaps the first 'democratising' technology, the printing press. Illich (1981) challenges the notion that the printing press eroded the

monopoly on knowledge held by institutions such as the Catholic Church, instead contending that printing accelerated the process of *standardising acceptable knowledge*, which in turn entrenched the power of already powerful institutions.

Taking these criticisms seriously, one may contend generative AI poses a risk to underrepresented groups when decision-makers oppose those groups. Specifically, there is a risk that generative AI becomes a *techno-solutionist* response to representation problems.

Techno-solutionism can be a broad and sometimes poorly defined term, often utilised in discussions seeking to criticise or dismiss different technological applications (Sætra and Selinger, 2024). As Sætra and Selinger (2024) note, the idea has its origins in an evolution-*cum*-critique of earlier notions that social problems can have technical solutions. Where technologies can solve social problems, the notion of a techno-solution should not be seen as a pejorative. For instance, two communities might fight over scarce water resources, but technology could reduce or remove this scarcity. Here, technology solves a social problem (fighting), and the technology could be labelled a techno-solution. Few would consider this techno-solution a negative. Morozov's (2013) use of the term techno-solution*ism*, however, is explicitly used as a normative term with negative connotations. Amongst other aspects, Morozov's techno-solutionism criticises the political deployment of technologies *to present the appearance of solving a social problem, without actually solving it*. It is in this normative sense that we refer to generative AI as techno-solutionist, in this particular instance.

One scenario to consider is that opponents of a particular group or perspective use generative AI to give the *appearance* of considering a plurality of views. Technically, generative AI produces outputs which *could* be used to consider many perspectives and tackle underrepresentation. But these outputs could also be used as 'evidence' that such views were considered, when in fact they never were. The effect of this would be comparable to Stiglitz's (2024) "façade of inclusiveness," where a seat at the table is used to disguise the fact that the microphone is, still, switched off.

Whether such a scenario would arise because of malicious intent, or circumstance, depends on the nature of the opposition. For instance, groups may be excluded, and decisions taken, wholly because of cost constraints. The applications of generative AI discussed in the algorithmic fidelity and AI innovation literatures do not suggest the technology can challenge common organisational cost constraints, but that it could offer ways for constrained decision-makers to act within those constraints. If decision-makers face pressure to *appear* to be considering a diversity of perspectives, without requisite organisational resources to actually do so, generative AI applications such as AI simulation may come to be used because they can achieve a given appearance within existing cost constraints. As above, cost, rather than accuracy, is the most cited benefit of AI simulation (Agnew et al., 2024).

Where generative AI is used to offer this appearance and lend apparent legitimacy to decision-makers for their own ends, the techno-solutionist aspect of the technology reveals itself as a more sinister force *in lieu* of perspectives advocated by Morozov (2011) and Illich (1981). Some experimental evidence may support this interpretation. For instance, one study of AI recommendations in policing decisions has found that police tend to follow such recommendations *only* when the recommendations align with what the individual *already* wanted to do (Selten et al., 2023). In such an instance, AI becomes a technological façade which gives the appearance of more 'objective' policing without challenging the decision-making structures which determine policing action.

In summary, perspectives are often omitted due to opposition by powerful interests and structural forces. Generative AI is a tool to generate information through simulation and prediction. But such information does not necessarily challenge these structural impediments to representation. Because those who voice opposition to underrepresented perspectives could themselves use generative AI outputs to promote the *appearance* of inclusivity; absent ways of challenging structural impediments to inclusion, generative AI may operate as a techno-solutionist 'solution' to representation, rather than a genuine one.

4.5. Temporal considerations

Rarely is a decision taken only once. This may be particularly true of sustainable development, where demands for sustainability must regularly be assessed against the impacts of development. Instead, decisions must often be taken many times, and (re-)evaluated over time (just as much as the consequences of decisions arise over time, too). Past decisions may establish the constraints upon which future decisions can be taken, thus impacting which perspectives and ideas can practically be pursued (Simon, 1997). Again, a key example of such a phenomenon is found in sustainable development, within discussions around "net zero" goals, which entail endeavouring to cut and offset greenhouse gas emissions to the degree that one's activities are carbon neutral (or even positive). Past development decisions now constrain the set of future development decisions which can be taken while maintaining net zero commitments.

The temporal aspect of representation is also relevant when considering generative AI. Generative AI technologies continue to develop, with measurable improvements in some aspects of performance frequently seen (at the time of writing). An immediate observation is that some of the limits of generative AI today, as a tool for representative decision-making, may be overcome in the future (AI Index Report, 2025). However, through considerations of time, one may approach the question of representation somewhat differently, and ask whether a generative AI model is *temporally representative*? This is to say, does a generative AI model have up-to-date information to, say, accurately simulate a population or formulate a policy recommendation?

Asking such a question leads to interesting observations which are likely to see greater discussion as generative AI becomes more integrated into organisational decision-making. For instance, the opportunities for generative AI to support greater representation may be limited by which AI system is selected for use in the decision-making process. OpenAI—creator of GPT-4—has emphasised the cutting-edge intelligence capabilities of their models (e.g., OpenAI, 2025). Anthropic, another generative AI company, emphasises their commitments to 'AI safety' and human values in their development of AI technologies (e.g., Anthropic, 2023). xAI, founded by the serial technology entrepreneur Elon Musk, has sought to differentiate their Grok AI model by emphasising commitments to freedom of speech and of political expression (e.g., Criddle and Murphy, 2025). Even at this nascent stage, the market for commercial AI products reflects how different values can influence design choices within these products, and these choices *may* influence how suitable models are when representing different perspectives.

Furthermore, when reflecting on decision-making over time, there is the possibility of an AI system becoming, in a manner of speaking, out-of-date. There are two perspectives this critique might take. Firstly, once a decision has been taken, the facts and actions concerning relevant groups is also likely to change. For instance, the objections of an indigenous group opposed to oil drilling are likely to evolve if, after a drilling license is granted, those people are then forcibly evicted from their land. Climate activist Roger Hallam (2020) has argued that environmental activists should strategically create uncertainty, through their actions, to undermine effective responses from those whom they politically oppose. This is to say nothing of the uncertainty decisions about sustainability might create once initiated, given the complexity of ecosystems, economies, and societies. Thus, in the abstract, decisions might cause the picture of reality possessed by an AI system to drift from the world as it actually is, leaving the system 'out-of-date' as a tool for decision-making and effective representation. The practical aspects of this 'knowledge drift' remain to be seen, and this is perhaps more of a speculation, but it is an interesting speculation, nevertheless.

One retort may be that AI systems, increasingly, have access to the internet, meaning they are able to present, and to an extent utilise, up-to-date information in their interactions with users. This functionality is likely to be important in forestalling 'knowledge drift' over time. However, on a technical level, up-to-date information would not be used by the AI system to update the weights within the system, and thus the 'learned' relationships found within natural language. These weights are derived from training data, and so a model trained in 2022 would respond to information about an event in 2025 in terms of the statistical relationships found in 2022.

The question of a model becoming out-of-date, therefore, is in part a question of whether language meaningfully changes within the timeframes of new models being trained? Companies like OpenAI and Anthropic frequently release new, updated models. However, there might be some instances—natural disasters, sudden scientific breakthroughs, political scandals—which prompt sudden cultural changes, and where even relatively new models may become out-of-date in terms of their weights. Furthermore, situating an AI system within an organisational context adds another dimension to this question. If organisations fail to shift to up-to-date models, say due to organisational constraints around cost, the AI systems which actually come to be used in decision-making may frequently be outdated compared to those which, in principle, could be available on the market.

This latter perspective speaks to a wider point about the temporality of AI systems. Updating an AI model to account for recent events may be a trivial problem compared to shifting social attitudes towards prejudiced groups or improving organisational procedures which resist change (Kingdon, 2010). The opportunities and risks as presented in Table 1 largely overlook that, in some instances, technologies can change much faster than people. As such, it is wise to attenuate the opportunities and risks outlined above with a recognition of the role of time in changing people, organisations, and technologies.

5. Section 5: recommendations and future research

The above analysis points to various recommendations which might be pursued by decision-makers and organisations engaged in sustainable development, and beyond.

Firstly, while generative AI might overcome some of the informational challenges associated with representative decision-making, there remain difficulties, many resulting from the problem of not knowing what one does not know. This leads to an immediate—perhaps obvious—consideration-*cum*-recommendation: involve underrepresented groups within the decision-making process. If the accuracy of a generative AI system remains in doubt or, from a temporal perspective, cannot adequately respond to changing decision environments; involving underrepresented groups would offer immediate recourse to this problem. Methods such as AI simulation or AI innovation can only be verified through comparison to the simulated group or consultation on the feasibility of a proposal, respectively. As such, one might conclude that generative AI acts as an unnecessary intermediary between stakeholders.

As above, the involvement of underrepresented groups—even if *normatively* preferable—may not be a practical solution to problems of underrepresentation in some instances. This may be because of individual opposition to involvement, or because of organisational constraints (*opposed, whether omitted or not*). Furthermore, underrepresented groups might themselves be inaccessible to decision-makers. For instance, some groups may lack the infrastructure or knowledge to engage with public calls for information, while some groups (say, anti-vaccine groups) might be outright hostile to decision-making bodies (say, public health bodies), despite the best efforts of the latter. Nevertheless, the most immediate solution to underrepresentation—namely, *engaging* underrepresented groups—should not be ignored.

Secondly, one might consider the development of generative AI models themselves. Much of the reviewed literature has focused on 'off-the-shelf' generative AI models, such as GPT-4. So too has much of the discussion in this article. However, it is conceivable that many of the advantages of generative AI technologies may be realised, without existing limitations arising from narrow training data, through locally developed or purpose-built generative AI systems. Organisations might choose to build their own generative AI systems so that innovation activities more closely align with practical, organisational constraints. Bespoke training data, such as in-depth interview data, might also be incorporated into a purpose-built model to more realistically simulate the perspectives of different peoples and groups. Where it is not practically feasible for all people to directly contribute to a decision, that group might choose to develop an AI model trained to represent the median preference of the group, rather than trusting their group representation to an individual who may prioritise their personal motivations. These activities respond to the problems of accuracy and feasibility discussed throughout this article and represent interesting lines of further research for future scholars.

Though, these proposals still raise questions of authority and organisational constraints. As above, one advantage of 'off-the-shelf' models is that they are immediately accessible to groups who have only limited resources, and that this accessibility could challenge asymmetries of expertise (Specian, 2023). If purpose-built models are needed, or at least offer advantages to those who have them, such asymmetries may remain. Furthermore, as we have sought to highlight in this article, omission is not just a product of ignorance but often follows from personal motivations and organisational constraints. This is likely to have substantial implications for the design of any purpose-built AI system. For instance, the data which are used in the training of the system will likely be influenced by the motivations of key decision-makers, while only data which are available or practically attainable could be used. If these factors already constrain representative decision-making, or work to the exclusion of some group; it may be reasonable to anticipate such exclusionary outcomes being recreated in the construction of a bespoke AI system. One recommendation for scholars and policymakers, given the problems of representation we have outlined in this article, is to undertake practical research into the suitability and acceptability of bespoke *versus* generic AI systems.

Finally, while this article has discussed generative AI, it has largely focused on LLMs. This is because LLMs are the most dominant kind of generative AI system within the literatures we have examined. Yet, generative AI captures many different kinds of AI systems with different modalities, such as image and audio generators (Sætra, 2023). These different kinds of generative AI system could raise some worthwhile questions around representation, specifically around ways of knowing and experiencing the world. For instance, there is some evidence that audio and visual experiences of indigenous environments positively influence conservation beliefs more than simply informing people about conservation issues (Banerjee and Ferreira, 2024). Generative AI, beyond text generators, may augment the ways decision-makers come to know and experience relevant perspectives. This is a dimension of representation we have not elaborated on here but is likely important in future studies of generative AI, sustainability, and representation. Therefore, a further recommendation, or possibility for future research, is exploring the intersection of representation and generative AI beyond text-generating systems.

6. Section 6: conclusion

This article has examined how generative AI can support more representative decision-making, with a specific focus on the implications for sustainable development and the sustainable development goals (SDGs). We have argued that perspectives can be overlooked because of the cognitive limitations of individual decision-makers (*omitted, but not opposed*). But perspectives can also be overlooked because of the opposition of key decision-makers, or the constraints faced by decision-

making bodies which make some perspectives conflictual or practically unfeasible (*opposed, whether omitted or not*).

Using these 'categories of omission' we have examined how generative AI might promote representative decision-making. We argue there are opportunities for promoting better representation which could arise from these applications, and that these opportunities could support the SDGs. When *omitted, but not opposed*, generative AI could highlight overlooked perspectives and interrogate pre-existing positions. When *opposed, whether omitted or not*, generative AI may improve access to expertise and support communication, empowering marginalised groups. There are also likely to be risks. We highlight problems related to ignorance (when perspectives are *omitted, but not opposed*) and techno-solutionism (when perspectives are *opposed, whether omitted or not*). We also note how temporality further complicates applications of generative AI, though contend that the analysis presented here represents a useful lodestar for future discussions, as the adoption of and experimentation with generative AI continues.

The relative novelty of generative AI means this discussion often draws upon hypotheticals or limited real-world examples, while the complexity of sustainable development inevitably means various perspectives could emerge supporting or opposing the arguments put forth here. However, by critically considering *why* decision-makers omit perspectives, we can speculate in a more structured, and thus practically useful, way, about the role of generative AI in representative decision-making.

CRediT authorship contribution statement

Stuart Mills: Writing – review & editing, Writing – original draft, Conceptualization. **Henrik Skaug Sætra:** Writing – review & editing.

Declaration of competing interest

The authors have no interests to declare.

Data availability

No data was used for the research described in the article.

References

Agnew, W.A., Bergman, S., Chien, J., Díaz, M., El-Sayed, S., Pittman, J., Mohamed, S., McKee, K.R., 2024. The illusion of artificial inclusion. Proceedings of the CHI'24 Conference on Human Factors in Computing Systems, pp. 1–12. https://doi.org/10.1145/3613904.3642703, 2024.

Aher, G., Arriaga, R.I., Kalai, A.T., 2023. Using large language models to simulate multiple humans and replicate human subject studies. In: Proceedings of the 40th International Conference on Machine LEARNING, 202, pp. 337–371.

AI Index Report, 2025. '*Artificial intelligence Index report 2025*' stanford human-centered artificial intelligence. Available at: https://hai.stanford.edu/assets/files/hai_ai_index_report_2025.pdf.

Alon-Barkat, S., Busuioc, M., 2023. Human-AI interactions in public sector decision making: "automation bias" and "selective adherence" to algorithmic advice. J. Publ. Adm. Res. Theor. 33, 153–169.

Amirova, A., Fteropoulli, T., Ahmed, N., Cowie, M.R., Leibo, J.Z., 2024. Framework-based qualitative analysis of free responses of large language models: algorithmic fidelity. PLoS One 19 (3), e0300024.

Anthropic, 2023. '*Core Views on AI safety: when, why, what, and how*' anthropic. Available at: https://www.anthropic.com/news/core-views-on-ai-safety.

Aonghusa, P.M., Michie, S., 2020. Artificial intelligence and behavioral science through the looking glass: challenges for real-world application. Ann. Behav. Med. 54 (12), 942–947.

Argyle, L.P., Busby, E.C., Fulda, N., Gubler, J.R., Rytting, C., Wingate, D., 2023. Out of one, many: using language models to simulate human samples. Polit. Anal. 31 (3), 337–351.

Banerjee, S., Ferreira, A., 2024. Virtual reality is only mildly effective in improving forest conservation behaviors. Sci. Rep. 14, e28532.

Bilgram, V., Laarmann, F., 2023. Accelerating innovation with generative AI: AI-augmented digital prototyping and innovation methods. IEEE Eng. Manag. Rev. 51 (2), 18–25.

Boden, M.A., 1998. Creativity and artificial intelligence. Artif. Intell. 103, 347–356.

Bouschery, S.G., Blazevic, V., Piller, F.T., 2023. Augmenting human innovation teams with artificial intelligence: exploring transformer-based language models. J. Prod. Innovat. Manag. 40, 139–153.

Brand, J., Israeli, A., Ngwe, D., 2024. 'Using GPT for market research' EC'. In: 24 Proceedings Of the 25th ACM Conference On Economics And Computation https://doi.org/10.1145/3670865.3673479.

Brem, A., Giones, F., Werle, M., 2023. The AI digital revolution in innovation: a conceptual framework of artificial intelligence technologies for the management of innovation. IEEE Trans. Eng. Manag. 70 (2), 770–776.

Cantens, T., 2025. How will the state think with ChatGPT? The challenges of generative artificial intelligence for public administrations. AI Soc. 40, 133–144.

Coeckelbergh, M., 2025. AI and epistemic agency: how AI influences belief revision and its normative implications. Soc. Epistemol. https://doi.org/10.1080/02691728.2025.2466164.

Cohen, M.D., March, J.G., Olsen, J.P., 1972. A garbage can model of organizational choice. Adm. Sci. Q. 17 (1), 1–25.

Criddle, C., Murphy, H., 2025. '*Elon Musk's Risqué Grok AI Chatbot Offers Challenge to Risk-Averse Rivals*. The Financial Times. Available at: https://www.ft.com/content/ca21ddb5-6391-476f-aace-ad181ab65762.

Cyert, R.M., March, J.G., 1992. A Behavioral Theory of the Firm, second ed. Prentice Hall, USA.

De Beukelaer, C., 2023. Trade Winds: A Voyage to a Sustainable Future for Shipping. Manchester University Press, UK.

Dunne, D., 2020. '*Mapped: How climate change disproportionately affects women's health*' Carbon Brief. Available at: https://www.carbonbrief.org/mapped-how-climate-change-disproportionately-affects-womens-health/.

EFRAG, 2024. '*EFRAG IG 1: materiality assessment*' EFRAG. Available at: https://www.efrag.org/sites/default/files/sites/webpublishing/SiteAssets/IG%201%20Materiality%20Assessment_final.pdf.

Feyerabend, P., 1978. Science in a Free Society. Verso Books, UK.

Füller, J., Hutter, K., Wahl, J., Bilgram, V., Tekic, Z., 2022. How AI revolutionizes innovation management—perceptions and implementation preferences of AI-based innovators. Technol. Forecast. Soc. Change 178, e121598.

Gmyrek, P., Lutz, C., Newlands, G., 2024. A technological construction of society: comparing GPT-4 and human respondents for occupational evaluation in the UK. Br. J. Ind. Relat 63 (1), 180–208.

Grove, A.S., 1997. Only the Paranoid Survive. Profile Books, UK.

Haefner, N., Wincent, J., Parida, V., Gassmann, O., 2021. Artificial intelligence and innovation management: a review, framework, and research agenda. Technol. Forecast. Soc. Change 162, e120392.

Hallam, R., 2020. Common sense for the 21st century: Only nonviolent rebellion can now stop climate breakdown and social collapse. Available at: https://rogerhallam.com/content/files/2023/12/Common-Sense-V0.5.1.pdf.

Hämäläinen, P., Tavast, M., Kunnari, A., 2023. Evaluating large language models in generating synthetic HCI research data: a case study. CHI' 23 (2023), 1–19. https://doi.org/10.1145/3544548.3580688.

Hardwig, J., 1985. Epistemic dependence. J. Philos. 82 (7), 335–349.

Herd, P., Moynihan, D., 2018. Administrative Burden: Decision-Making by Other Means. Russell Sage Foundation, USA.

Hoefer, R., 2022. The multiple streams framework: understanding and applying the problems, policies, and politics approach. J. Pol. Pract. Res. 3, 1–5.

Howard, T.J., Culley, S.J., Dekoninck, E., 2008. Describing the creative design process by the integration of engineering design and cognitive psychology literature. Des. Stud. 29 (2), 160–180.

Howarth, C., Parsons, L., Thew, H., 2020. Effectively communicating climate science beyond academia: harnessing the heterogeneity of climate knowledge. One Earth 2 (4), 320–324.

Hwang, E., Majumder, B. P., Tandon, N. (2023) 'Aligning Language Models to User Opinions' Available at: https://arxiv.org/abs/2305.14929.

Illich, I., 1981. *Shadow Work*' Marion Boyars: USA.

Kakatkar, C., Bilgram, V., Füller, J., 2020. Innovation analytics: leveraging artificial intelligence in the innovation process. Bus. Horiz. 63, 171–181.

Kingdon, J.W., 2010. Agendas, Alternatives, and Public Policies. Pearson, USA.

Lakhani, N., 2021. "*A continuation of colonialism': indigenous activists say their voices are missing at Cop26'* the Guardian. Available at: https://www.theguardian.com/environment/2021/nov/02/cop26-indigenous-activists-climate-crisis .

Landemore, H., 2022. Open Democracy: Reinventing Popular Rule for the Twenty-First Century. Princeton University Press, UK.

Lee, S., Peng, T., Goldberg, M.H., Rosenthal, S.A., Kotcher, J.E., Maibach, E.W., Leiserowitz, A., 2024. Can large language models capture public opinion about global warming? An empirical assessment of algorithmic fidelity and bias. PLOS Climate 3 (8), e0000429.

Lindblom, C.E., 1959. The science of muddling through. Public Adm. Rev. 19 (2), 79–88.

Maas, A.H.E.M., Appelman, Y.E.A., 2010. Gender differences in coronary heart disease. Neth. Heart J. 18, 598–603.

March, J.G., Simon, H.A., 1993. Organizations, second ed. Wiley, USA.

Mariani, M.M., Machado, I., Magrelli, V., Dwivedi, Y.K., 2023. Artificial intelligence in innovation research: a systematic review, conceptual framework, and future research directions. Technovation 112, e102623.

Marx, P., 2022. Road to Nowhere: what Silicon Valley Gets Wrong about the Future of Transportation. Verso Books, UK.

Mei, Q., Xie, Y., Yuari, W., Jackson, M.O., 2024. A Turing test of whether AI chatbots are behaviorally similar to humans. Proc. Natl. Acad. Sci. 121 (9), e2313925121.

Merone, L., Tsey, K., Russell, D., Nagle, C., 2022. Sex inequalities in medical research: a systematic scoping review of the literature. Women's Health Rep. 3 (1), 49–59.

Mohamed, J., Anderson, P., Matthews, V., 2022. Indigenous peoples across the globe are uniquely equipped to deal with the climate crisis—so why are we being left out of these conversations? The Conversation. Available at: https://theconversation.com/indigenous-peoples-across-the-globe-are-uniquely-equipped-to-deal-with-the-climate-crisis-so-why-are-we-being-left-out-of-these-conversations-171724.
Moore, B., Geese, L., Kenny, J., Dudley, H., Jordan, A., Pascual, A.P., Lorenzoni, I., Schaub, S., Enguer, J., Tosun, J., 2024. Politicians and climate change: a systematic review of the literature. WIREs Clim. Change 15 (6), e908.
Morozov, E., 2013. To Save Everything, Click Here: Technology, Solutionism, and the Urge to Fix Problems that Don't Exist. Allen Lane, UK.
Morozov, E., 2011. The Net Delusion: How Not to Liberate the. *World*' Penguin Books, UK.
Neville, S., 2024. '*How medical research is failing women*' the Financial Times. Available at: https://www.ft.com/content/ce9895f9-8ead-4b43-b473-874aebabc0e6.
OpenAI, 2025. '*Introducing OpenAI o3 and o4-mini*' OpenAI. Available at: https://openai.com/index/introducing-o3-and-o4-mini/.
Oreskes, N., Conway, E.M., 2010. Defeating the merchants of doubt. Nature 465, 686–687.
Ostrom, E., 1996. Crossing the great divide: coproduction, synergy, and development. World Dev. 24 (6), 1073–1087.
Ostrom, E., 1990. Governing the Commons: the Evolution of Institutions for Collective Action. Cambridge University Press, UK.
Peterson, A.J., 2025. AI and the problem of knowledge collapse. AI Soc. https://doi.org/10.1007/s00146-024-02173-x.
Pettifor, A., 2019. The Case for the Green New Deal. Verso Books, UK.
Polanyi, M., 2005. Personal Knowledge: towards a Post-Critical Philosophy. Routledge, UK.
Santurkar, S., Durmus, E., Ladhak, F., Lee, C., Liang, P., Hashimoto, T., 2023. Whose opinions do language models reflect? Proceedings Of the 40th International Conference On Machine Learning.
Sætra, H.S., 2024b. '*Computation and deliberation: the Ghost in the habermas machine*' ACM communications blog. Available at: https://cacm.acm.org/blogcacm/computation-and-deliberation-the-ghost-in-the-habermas-machine/.
Sætra, H.S., 2024a. '*The ESG backlash: politics, ideology, and the Future of sustainable business*' SSRN. Available at: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4782018.
Sætra, H.S., 2023. Generative AI: here to stay, but for good? Technol. Soc. 75, e102372.
Sætra, H.S., 2022. AI for the Sustainable Development Goals. CRC Press, UK.
Sætra, H.S., 2020. A shallow defence of a technocracy of artificial intelligence: examining the political harms of algorithmic governance in the domain of government. Technol. Soc. 62, e101283.
Sætra, H.S., Selinger, E., 2024. Technological remedies for social problems: defining and demarcating techno-fixes and techno-solutionism. Sci. Eng. Ethics 30 (60), 1–17.
Sætra, H.S., Selinger, E., Mills, S., 2025. Artificial intelligence (and conflict). In: Samoilenko, S.A., Simmons, S. (Eds.), The Handbook of Social and Political Conflict. Wiley, USA, 2025.
Schoenbegger, P., Tuminauskaite, I., Park, P.S., Tetlock, P.E., 2024. Wisdom of the silicon crowd: LLM ensemble prediction capabilities rival human crowd accuracy. Sci. Adv. 10 (45), e1528.
Selten, F., Robeer, M., Grimmelikhuijsen, S., 2023. Just like I thought: street-level bureaucrats trust AI recommendations if they confirm their professional judgment. Public Adm. Rev. https://doi.org/10.1111/puar.13602.
Sen, A., 2002. Rationality and Freedom. Harvard University Press, USA.
Shreedhar, G., Moran, C., Mills, S., 2024. Sticky brown sludge everywhere: can sludge explain barriers to green behaviour? Behaviour. Pub. Pol. 1–16. https://doi.org/10.1017/bpp.2024.3.
Shrestha, P., Krpan, D., Koaik, F., Schnider, R., Sayess, D., Binbaz, M.S., 2025. Beyond WEIRD: unveiling GPT's cultural bias in global policy research? Behavior. Sci. Pol. https://doi.org/10.1177/23794607241311793.
Si, C., Yang, D., Hashimoto, T., 2024. Can LLMs generate novel research ideas? A large-scale human study with 100+ NLP researchers. Arxiv. Available at: https://arxiv.org/html/2409.04109.
Simon, H.A., 1997. Administrative Behavior, fourth ed. Free Press, USA.
Simon, H.A., 1987a. Two heads are better than one: the collaboration between AI and OR. Interfaces 17 (4), 8–15.
Simon, H.A., 1987b. Making management decisions: the role of intuition and emotion. Acad. Manag. Exec. 1 (1), 57–64.
Simon, H.A., 1981. The Sciences of the Artificial. MIT Press, USA.
Simon, H.A., 1956. Rational choice and the structure of the environment. Psychol. Rev. 63 (2), 129–138.
Simon, H.A., 1955. A behavioral model of rational choice. Q. J. Econ. 69, 99–118.
Sorensen, T., Moore, J., Fisher, J., Gordon, M., Mireshghallah, N., Rytting, C.M., Ye, A., Jiang, L., Lu, X., Dziri, N., Althoff, T., Choi, Y., 2024. A roadmap to pluralistic alignment. In: Proceedings of the 41st International Conference on Machine Learning, 235, pp. 46280–46302.
Specian, P., 2024. Machine advisors: integrating large language models into democratic assemblies. Soc. Epistemol. https://doi.org/10.1080/02691728.2024.2379271.
Specian, P., 2023. '*Large Language Models for democracy: Limits and possibilities*' technology and sustainable development 2023 conference paper. Available at: https://techandsd.com/_files/specian_2023.pdf.
Stallard, E., 2022. COP27: lack of women at negotiations raises concern. BBC News. Available at: https://www.bbc.co.uk/news/science-environment-63636435.
Stiglitz, J., 2024. The Road to Freedom. Allen Lane, UK.
Sue, D.W., 2006. The invisible whiteness of being: whiteness, white supremacy, white privilege, and racism. In: Constantine, M.G., Sue, D.W. (Eds.), Addressing Racism: Facilitating Cultural Competence in Mental Health and Educational Settings. Wiley, USA, 2006.
Tessler, M.H., Bakker, M.A., Jarrett, D., Sheahan, H., Chadwick, M.J., Koster, R., Evans, G., Campbell-Gillingham, L., Collins, T., Parkes, D.C., Botvinick, M., Summerfield, C., 2024. AI can help humans find common ground in democratic deliberation. Science 386 (6719), e2852.
Tversky, A., Kahneman, D., 1974. Judgment under uncertainty: heuristics and biases. Science 185 (4157), 1124–1131.
Tversky, A., Kahneman, D., 1973. Availability: a heuristic for judging frequency and probability. Cogn. Psychol. 5 (2), 207–232.
United Nations, 2015. Transforming Our World: the 2030 Agenda for Sustainable Development. ' Division for Sustainable Development Goals. Available at: https://sustainabledevelopment.un.org/content/documents/21252030%20Agenda%20for%20Sustainable%20Development%20web.pdf?ref.
Wouter Botzen, W.J., Thepaut, L.D., Banerjee, S., 2025. Kahneman's insights for climate risks: lessons from bounded rationality, heuristics and biases. Environ. Resour. Econ. https://doi.org/10.1007/s10640-025-00980-4.