ARTICLE

Check for updates

Can digital tools foster ethical deliberation?

Joanna Sleigh [1✉], Shannon Hubbs[1], Alessandro Blasimme [1] & Effy Vayena[1]

In the age of Machine Learning Algorithms, Artificial Intelligence and Natural Language Processing, digital technologies have become interwoven in various aspects of society, including in our practices for ethical deliberation and decision-making. In this study, we present a systematic mapping and taxonomy of digital tools designed explicitly for this purpose and published between 2010 and 2023. By providing a comprehensive overview of the landscape, we identify the key features and mechanisms employed to facilitate ethical deliberation. This research enhances our understanding of the potential role that digital tools can play in supporting ethical decision-making processes, offering valuable insights for ethicists, educators, government organizations, and private institutions seeking to develop, deploy, or utilize such tools.

[1] Health Ethics and Policy Lab, Department of Health Science and Technology, ETH Zürich, Switzerland. ✉email: joanna.sleigh@hest.ethz.ch

Introduction

In today's media and technology landscape, numerous digital tools have emerged to tackle 21st century ethical issues and challenges. Recent examples include Microsoft's "Ethical Artificial Intelligence Toolbox" and the MIT Media Lab's "Moral Machine," which grapple with ethical dilemmas in artificial intelligence (AI) and machine learning (ML) (Awad et al., 2018; Wong et al., 2023). Yet amidst the proliferation of such technologies, the domain of tools designed to aid humans in ethical deliberation remains unexplored. This paper delves into this area, offering an analysis and comparative examination of computerised tools made to support end-users in their ethical deliberation. The emphasis lies in dissecting the technological mechanisms and validation approaches of the tools and considering the potential advantages or drawbacks for both individuals and society that arise from their utilisation. In doing so, this paper serves as a resource for ethicists, educators, government organisations, and private institutions interested in the development and use of ethical deliberation tools online.

Background. Ethical deliberation refers to an agent's ability to "discuss openly and reflect on understandings of moral problems, on solutions to these problems, and to explore what a meaningful resolution could amount to…" (Senghor and Racine, 2022, p. 1). Drawing on Dewey's pragmatist ethics, this process comprises three key moments: the recognition of a morally problematic situation, the imagination of different scenarios, and the evaluation of various solutions to inform a judgment or action (Senghor and Racine, 2022). This conceptualisation aligns with Aristotle's notion of ethical deliberation, whereby a moral agent internally deliberates and continues to search for a resolution (Goodin and Niemeyer, 2003). Scholars add that ethical deliberation's goal is not necessarily a definitive conclusion, but rather "enrichment of one's own point of view with that of others" (Gracia 2003, p. 227). Indeed, stemming from philosophy and the political sciences, deliberation (*bouleusis*) can be an internal reflection or a collective discussion that takes into consideration the perspectives of impacted stakeholders (Gutmann and Thompson, 1997; Goodin and Niemeyer, 2003; Ten Have and Patrão Neves, 2021). Based on this theoretical framework, we henceforth understand ethical deliberation as a reflective process that can be individual or collective, and one that while being potentially conducive to ethical decision-making, may or may not result in a definitive decision by the user.

To facilitate ethical deliberation, a plethora of digital and analogue tools are available, encompassing methodologies, frameworks, processes, guidelines, matrices, and codecs (Beekman and Brom, 2007). Their common characteristics, as described by Moula and Sandin (2015, p. 264), are that they are designed for versatility, accommodate diverse ethical viewpoints, and function as heuristic aids rather than rigid decision-making algorithms. Based on a review of 60 methods, Maner (2002) outlines twelve stages present in ethical deliberation tools, from the cultivation of moral awareness to monitoring and implementing a decision. Laaksoharju (2010) cautions that employing all twelve stages may prove cumbersome in practice and recommends tools rather be defined by how they seek to bolster human autonomy, specifically an individual's capacity to reason independently. Accordingly, tools serving only informational purposes or promoting algorithmic decision making cannot be classified as deliberative. This category includes tools suited purely for information-gathering, such as simulations (Diallo et al., 2021) and tools that automatically evaluate ethical issues based on the software designer's ethical interpretation (Zhuo et al., 2023; Zuber et al., 2022; Brown and Mecklenburg, 2021;

Zhang-Kennedy and Chiasson, 2021). In summary, we define ethical deliberation tools as inherently human-centric, offering users systematic and structured approaches to navigate ethical issues in their complexity, allowing exploration of various ethical options on their own merits and potentially supporting ethical decision-making.

However, defining the boundaries of ethical deliberation tools poses a challenge, given that many technologies lack explicit 'ethical' labelling (Moula and Sandin, 2015). Consider how online games and visual novels, such as Bioshock (Travis, 2010) and Eliza (Ramos, 2019), despite not being explicitly designed for ethics, can foster critical reflection on technological innovations. Research supports this perspective, indicating that gamification can enhance presence, emotional engagement, cognitive absorption, and ethical insight (Lyreskog et al., 2023). Ethical debates also extend to the use of persuasion and emotional appeals in these tools (Kim and Werbach, 2016). A parallel discourse surrounds nudging techniques in tools like ethical shopping apps or investment software, with scholars weighing potential infringements on autonomy against contributions to users' digital self-control and well-being (Ienca and Vayena, 2021; Monge Roffarello and De Russis, 2023). This discussion is further supported by research that reveals technology-dependent variations in moral responses (Pan and Slater, 2011) and that emotions can impact moral decision-making (Navarrete et al., 2012). These findings raise questions about the ways in which different mechanisms can enhance or support ethical deliberation.

Consequently, how can we ascertain the quality of an ethical deliberation tool and gain insight into its mechanisms and impact on users? Kaiser et al. (2007) proposes that tools ought to be assessed for their ethical soundness, referring to a tool's alignment with ethical principles or theories. Other scholars recommend assessing the deliberation process itself (Towne and Herbsleb, 2012), examining criteria relating to ethical reflection or the imagination of scenarios (Senghor and Racine, 2022), along with usability, user-friendliness, user satisfaction, and ease of use (Xenos and Velli 2018; Marti and Iacono 2016). In contrast, Moula and Sandin (2015) advocate for evaluating a tool based on its intended outcome, whether it be to aid users with decision-making, or rather heuristic purposes. As an example, Stark et al. (2021) evaluated the outcome of preference transformation. Three key approaches stand out in this evaluation literature, namely measuring: a) the tool's ethical soundness; b) the quality of deliberation; c) and the resulting outcomes.

Despite this growing body of literature exploring individual ethics tools and methods of evaluation, no comprehensive overview of digital tools for ethical deliberation exists. This study aims to fill this gap by mapping the landscape of computerised tools for ethical deliberation. Our research questions are as follows:

R1: What mechanisms (i.e., checklists or scenarios) are used by digital ethics tools to promote ethical deliberation?

R2: Do these digital ethics tools provide evidence of effectiveness, specifically in terms of ethical soundness, quality of the ethical deliberation process, or achievement of intended outcomes?

Method

We conducted a systematic mapping review of online ethical tools published between 2010 and 2023. Systematic mapping, distinct from systematic review, employs broader inclusion criteria and offers a comprehensive overview of a specific field by "collating, describing, and cataloguing evidence" (James, Randall, and Haddaway, 2016, p. 1). This method was chosen for its capacity to

analyse and compare clusters of digital tools reported across diverse data sources, as evident in the works of various scholars (Zohud and Zein, 2019; Wimalasooriya et al., 2022; Mystakidis et al., 2022). Petersen et al.'s (2008) guidelines add that systematic mappings help identify research gaps in a specific topic area and indicate the absence of evaluation or validation research.

Our protocol for conducting the systematic mapping review encompassed six distinct stages. First, we defined the scope and formulated precise research questions. Second, we executed searches using a predefined strategy. Third, we screened papers for eligibility. Fourth, we incorporated coding and faceted analysis (Kwasnik, 1999). Fifth, we carried out a critical appraisal, investigating the overall validity of the evidence base and subsets of evidence. Finally, we engaged in the description, visualisation, and reporting of our findings.

Information sources and search strategy. The search strategy was informed by the protocols for data collection and identification of web-based resources proposed by Godin et al. (2015); Blasimme et al. (2018) and Jobin et al. (2019). Following best practices for grey literature search, and to achieve comprehensiveness, we incorporated several complementary sources: (1) grey literature databases, (2) targeted digital libraries, (3) customised Google search engines, and (4) expert consultation. First, we performed a grey literature search using the databases of PubMed, Scopus, and IEE Xplore and included the following keywords in the search strings: 'digital', 'ethics', 'decision', 'deliberation', and 'tool'. Due to the infrequent usage of the word 'deliberation' in non-academic settings, the term 'decision' was included to gain a wider sample of results. In complement, we searched five digital libraries (IOS App store, Google Play Store, Chrome web store, Microsoft Edge Add-ons, and GitHub). We further conducted a keyword-based search using the same terms on Google.com. Private-browsing mode was used, with web cookies and history deleted. Finally, we consulted with ethics experts to promote data saturation. Excel was used to document the identified contacts and the recommended tools and resources. Figure 1 presents the flow and numbers of selection, inclusion, and exclusion.

Eligibility criteria and classification. To assess eligibility for inclusion, we built a classification scheme based on three factors: a) ethical focus, b) digital format, and c) accessibility. As Table 1 shows, to qualify for inclusion, a tool had to be purposefully designed to promote ethical deliberation, defined as the ability to "discuss openly and reflect on understandings of moral problems, on solutions to these problems, and to explore what a meaningful resolution could amount to is highly valued for clear reasons" (Senghor and Racine, 2022, p. 1). We consequently excluded tools serving only informative functions (such as collections of educational resources, clinical decision support tools or patient information tools), along with tools designed only for computerised assessment (i.e., bias detection, ethical hacking, or technological assessment). Also excluded were tools that promoted debate and argumentation without an ethics focus (i.e., Kialo.com, and Debatemap.app). Second, included tools needed to have a digital or electronic format, harness digital qualities, and extend beyond a practical method or conceptual framework published on a website. As such, we excluded some digital formats, such as PowerPoint presentations, Portable Document Formats (PDFs), E-Books, data libraries, code packages and images (i.e., navigation wheels (Kvalnes and Kvalnes, 2019), or The Open Ethics Canvas (Lukianets et al., 2021)). Finally, we excluded tools that were no longer accessible, incomplete (i.e., prototypes proposed in academic papers), required payment, or were written in a language other than English.

Content analysis and taxonomy development. This study utilised a directed content analysis approach to describe the characteristics and mechanisms of ethical deliberation tools (Hsieh and Shannon, 2005). The coding process first involved categorising tools based on three moments of ethical deliberation (Senghor and Racine, 2022). We then inductively coded descriptive qualities such as technology type, tool author, publication date, intended target audience, general topic area, and whether the tool was designed for individual or group use. The inclusion of these descriptive elements was grounded in the recognition that the ethical deliberation facilitated involves more than just the technology itself, but encompasses a convergence of technical, political, and other decision-making factors, along with the
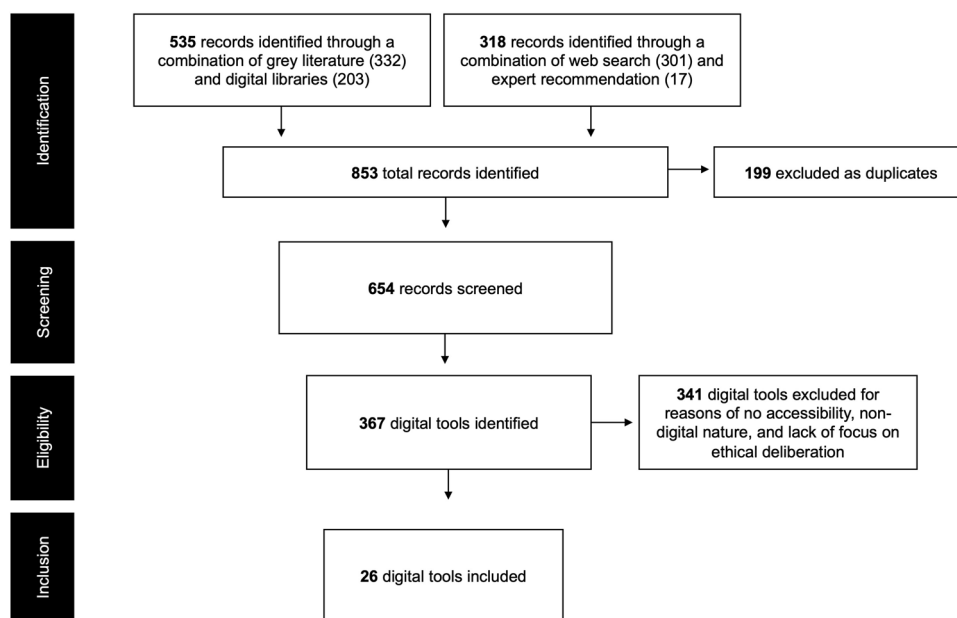

Fig. 1 Flowchart of tool selection.

Table 1 Details of inclusion and exclusion criteria.

Inclusion criteria	Exclusion criteria
The tool was developed with the intent of facilitating **ethical deliberation** and the user is part of the deliberation process.	The tool does not relate directly to ethics. The tool was not designed with the intention to be used for ethical deliberation (i.e., it is a commercial game, charity, or a fashion brand). The tool's content or descriptive data do not mention ethical reflection, decision making or deliberation. The tool does not involve the user in the decision-making / deliberation process or is elevated to an authority above the user. For example, it is a software or algorithm that makes autonomous decisions of what is right or wrong. The tool lacks a focus on moral and ethical reflection. It either offers data for decision making, as with technology assessment tools, or enforces predetermined ethical standards without encouraging critical reflection on these values.
The tool has a **digital format**, meaning a program, app, or software with a user interface and interactive capacity.	The tool's format is not digital or computerised. For example, it is only a practical method and/or conceptual framework published on a website (aka a PDF). The tool lacks digital qualities, having no user interface, being a static image, or serving only as a data collection tool.
The tool is **accessible** online or via download as an app, program, or software, and is available in English.	The tool is not open access. The tool is not available online, because it is a conceptualisation, has not been completed, is a prototype, or is no longer available. The tool is not available in English.

contextual contingencies in which these tools are employed (Wright and Street, 2007). Next, we inductively coded ethical deliberation mechanisms, referring to the features each tool utilised to facilitate ethical discussion, analysis, and exploration of different perspectives and options (see Table 2 for details).

To answer the research question asking whether the tools measured their impact or provided evidence of effectiveness in achieving their intended purposes or desired outcomes, a three-phase process was employed. First, a comprehensive content analysis was conducted to identify any reference to evaluations within the tool. Second, an online search was conducted using Google Scholar, utilising the private Incognito mode. Last, proactive outreach was undertaken to contact the developers of the tools and request additional information. Throughout the coding process, disagreements between the two researchers were discussed, and refinements were made to the coding categories until an agreement was reached.

Results

Characteristics of the analysed tools. Following the inclusion and exclusion criteria, the final sample was made up of 26 tools published between 2010–2023 (see Table 3). Analysis of their characteristics revealed variations in technology type, authors, audiences, publication dates and topic domains. As visualised in Figs. 2 and 3, most tools were web-based ($n = 21$, 81%), intended for individual use ($n = 19$, 73%), and published in 2021. Regarding the tool creators, most were authored by universities ($n = 14$, 54%), followed by the European Union's Horizon 2020 program ($n = 4$, 15%). The main intended audiences were developers ($n = 8$, 31%), academics ($n = 6$, 23%) and broad audiences ($n = 5$, 19%). Five topic domains were present in the analysed tools, mostly data usage ($n = 13$, 50%) and technology development ($n = 11$, 42%), followed by philosophy ($n = 11$, 42%), research integrity ($n = 7$, 27%), and health ($n = 5$, 19%).

Deliberative mechanisms. Analysis revealed the tools used a diverse array of mechanisms to facilitate ethical deliberation, encompassing both traditional elements alongside technology-driven techniques, with descriptions provided by Table 2. Traditional elements included scenarios, case studies, frameworks, checklists, supplementary resources, and question prompts. Digital features included interactive visualisation, online gamification, discussion

forums, and feedback mechanisms. All tools were found to combine deliberation mechanisms to help structure critical analysis, encourage reflective thinking, and deepen the level of engagement, whilst some even enabled personalisation and data collection for research.

Question prompts were the most frequently used deliberation mechanism ($n = 26$, 100%). These prompts played a dual role, serving to both stimulate and structure reflection, actively involving users in ethical deliberation. For instance, in the RRI Self-reflection tool, users must answer various questions about their research within the framework of research integrity. This tool allows users to save their responses and generate a final report, facilitating sharing with peers, ethics committees, or simply documenting and rendering transparent the deliberative process. Similarly, the Fairness Compass uses question prompts with feedback mechanisms to support the user in selecting and articulation the appropriate fairness definition for their AI system with broader goals to cultivate societal trust.

Visualisations emerged as the second most widely employed mechanism for deliberation ($n = 19$, 73%). For example, the Ethical Stack tool employed a layered visualisation approach to encourage multidisciplinary teams (comprising developers, designers, product managers, CEOs, and others) to systematically deconstruct their product and link dimensions such as data usage, 3rd party access, or context, to their ethical values. Conversely, the Trolly Game, a web-based interactive tool, depicted the classic "trolley dilemma" using black and white sketch-like illustrations. Users get presented with the decision of whether to pull a lever to save one life at the expense of another. While the former tool used visualisation to structure and clarify ethical dimensions and variables, the latter harnessed this mechanism to drive engagement, enhance accessibility, and enrich the storytelling component.

Feedback mechanisms ($n = 14$), resources ($n = 14$), scenarios ($n = 11$), and gamification ($n = 6$) were also frequently used by the tools analysed. One example which combined all four features was the Dilemma Game, a mobile app that presents users with research integrity scenarios. Users select which out of a series of potential solutions they agree with, and after submitting their choice, the tool reveals the percentage of agreement or disagreement amongst other participants. Additionally, the tool provides an ethics expert opinion with links to guidelines, as a resource for each case. Gamification is present when a user engages in the "group" setting, in which each participant must cast their vote on a dilemma. Once everyone has made their selection, each user's choice becomes visible, and they

Table 2 Codebook of categories and variables examined: moment of ethical deliberation, mechanism for ethical deliberation, and focus of the tool's evaluation.

Dimension	Category	Description
Moment of deliberation	Recognise and understand the issue	Ethical deliberation begins with recognition of an ethical issue. It then involves a deepening of understanding, by collecting facts and information, and considering different perspectives and values.
	Imagine or identify options to act	The second stage involves ideating or identifying possible scenarios to overcome the problem. This reflective phase takes into consideration the environment, people, actions, and context.
	Reflect and assess these options to come to a judgement	The third stage involves weighing the possible scenarios based on evaluation of their respective feasibility and ethical merits. This involves looking at their ethical acceptability (i.e., reduction of harms) and praiseworthiness (i.e., promotion of human flourishing).
Mechanism for deliberation	Question prompts	Questions guide discussions, foster critical thinking, and ensure that participants consider a wide range of perspectives and factors. They also promote a structured approach to ethical decision making.
	Visualisation	Visualisation refers to the use of graphs, charts, or infographics that represent information. Often interactive, they help users explore and identify patterns or trends in data. It does not refer to graphic design.
	Resources and information	Resources, such as the provision of assessment tools, research papers, videos, or expert opinions, can provide additional information or context relevant to the decision. They can also support exploration of different perspectives, options, or evidence.
	Scenarios and case studies	Scenarios describe possible situations or events. They help users imagine and consider different hypothetical outcomes or futures. Thereby helping users weigh and compare potential implications.
	Feedback mechanisms	Feedback mechanisms like surveys, ratings, or quantitative assessments, can provide data-driven results to help analyse and understand ethical issues and potential impacts to stakeholders.
	Checklist	Checklists are lists of items for users to systematically consider and address ethical concerns. They provide a structured approach to guide users through considerations in their actions or decisions.
	Gamification	We use the term gamification to broadly refer to the inclusion of game like elements such as points, badges, leader boards, and storytelling. These features encourage active participation and foster deeper engagement.
	Discussion Forum	Discussion forums are digital platforms where participants engage in dialogue, share ideas, and explore diverse perspectives, facilitating consensus-building and highlighting agreement or disagreement.
	Artificial Intelligence (AI)	AI refers to computer systems capable of learning from data and making informed judgments. It can help users analyse complex ethical dilemmas from diverse perspectives and assist in decision-making processes.
Focus of the tool's evaluation	Ethical soundness	Assessment of the extent to which the tool incorporates, and references established ethical theories, principles, or frameworks in its guidance.
	Process of deliberation	Examination of the tools' process of deliberation, focusing on the user experience and the effectiveness of the reflective process it enables.
	Outcome of deliberation	Investigation of the tool's ability to bring about changes in the user's learning and ethical reasoning. Focus is on the impact of the tool on the user's ethical decision-making capabilities.

are required to defend their decision. The tool thus uses a combination of mechanisms to engage users in collective or individual reflection around resolving ethical issues within a research context.

In one tool we also observed the use of AI as an ethical deliberation mechanism. The EDEN (Ethical Dilemma Evaluation Network) tool employs AI to create multiple chatbots, each representing a distinct ethical perspective. Underpinned by Python and driven by OpenAi's GPT-4 language model, the tool invites users to propose an ethical issue or dilemma and receive detailed answers that deconstruct the prompted issue based on the normative values and principles of different ethical theories. The user can then engage in weighing and comparing the different approaches.

Evidence of effectiveness. Analysis of the tools revealed a diversity of approaches for validating effectiveness and bolstering credibility. The predominant approach ($n = 22$, 85%) was normative, with tools grounding their design in ethical theories, principles, or frameworks. For instance, the Felicific Calculator claims to operationalise Jeremy Bentham's utilitarian philosophy, prompting users to assess questions regarding the pain and pleasure experienced by affected agents.

Alternatively, fifteen tools (58%) substantiated their effectiveness through peer-reviewed publications. Some publications scrutinised the quality of the deliberation process (10 tools), while others (8 tools) demonstrated the tool's impact, such as heightened ethical awareness or sensitivity. Certain tools had papers examining both aspects. Sleigh et al. (2023), for example, assessed their ethical deliberation tool's user experience and its influence on users' understanding, acquisition, application, analysis, and synthesis of ethical knowledge, informed by Bloom's (1956) and Krathwohl's (2002) theories. Similarly, the Ethxpert tool underwent evaluation through various mixed methods studies, gauging user satisfaction and the quality of ethical analysis. Notably, although it helped users comprehend an ethical

**Table 3 Names and descriptions of tools analysed.**

Tool	Description
Ethxpert	Software tool that aims to support a systematic, self-critical, and holistic approach to ethical decision making with no reference to any ethical theory.
Quandary	Web-based game for teaching ethical decision making and critical thinking skills to players of all ages.
Ethical Decision Making	App for helping individuals think through ethical issue by identifying stakeholders, ethical perspectives, and weighing options.
RRI Self-reflection Tool	Online tool for reflecting on research integrity and documenting answers.
DEDA for public management	Interactive questionnaire to raise awareness around using data ethically.
CDT so you want to code an algorithm	Visual tool with questions that guides ethical reflection around building algorithms.
Societal Readiness Thinking Tool	Online tool with questions to guide reflection around responsible research and innovation.
Recode: Digital Health Checklist	Decision support checklist to guide ethical digital health research.
Ethical systems map	Interactive visualisation of principles to foster reflection on ethical health research and data usage.
Ethix	Interactive questionnaire for identifying values and ethical risks.
Ethos 2514	App-based game to guide nuanced ethical discussions on dilemmas.
ML Ethics Tool	Web tool for identifying, understanding, and mitigating ethical concerns in machine learning.
Dilemma Game	App-based game to stimulate awareness and discussion about research integrity.
Ethical Stack	An interactive visualisation tool to structure and facilitate considering ethics when building new connected technology.
Felicific calculator	Calculator based on Felicific Calculus, translating the philosophy of Jeremy Bentham by quantifying pain and pleasure experienced by impacted agents.
My Bioethics	App focusing on healthcare ethics and bioethics.
FCAI Ethics Exercise Tool	Interactive questions for reflection at various program/project phases to help AI researchers consider the ethical implications of their research and strengthen ethical understanding.
Fairness compass	Interactive visual tool for users to select which fairness definition fits their system / AI applications, with documentation of the decision-making process seeking to promote explainability and broader societal trust.
Digital ethics compass	Interactive tool to guide companies in ethical design decisions in data and AI ethics.
Integrity Game	Web-based platform with scenarios for teaching academic integrity.
Clinical ethics	Structured app to help clinicians navigate decisions in situations of ethical uncertainty in clinical practice.
Pervada Data Ethics Support Tool	Tool using interactive questions to focus on health data ethics.
EPFL Game	Game-based approach to learn about ethics and potential outcomes of Machine Learning applications.
Interactive AI Ethics Quiz	Interactive quiz exploring AI developers' perception and approaches to ethics implementation to help overall ethics practice.
EDEN (Ethical Dilemma Evaluation Network)	Python-based program simulating ethical decision-making between AI-powered chatbots, each representing a distinct ethical perspective.
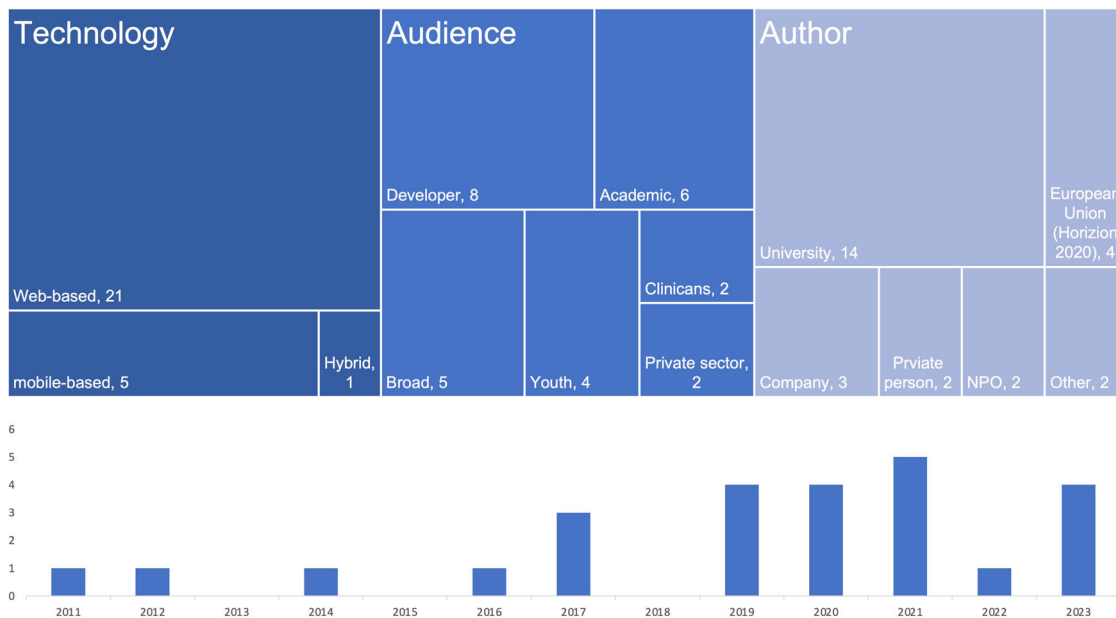
Fig. 2 Tree-map of the digital ethical tools' technology, audience, and authors, with a bar graph showing the dates of publications.

issue, it fell short in helping users come to a decision on what action to take (Laaksoharju and Kavathatzopoulos, 2002). Conversely, the Quandary tool employed mixed-method studies to evaluate its effectiveness in fostering ethical decision-making among school students (Hilliard et al., 2018; Ilten-Gee and Hilliard, 2021; Lawrence and Sherry, 2021). Results indicated significant improvements in fact vs. opinion comprehension, perspective-taking, teacher satisfaction and student engagement, underscoring Quandary's value as an educational tool for nurturing critical thinking and empathy.
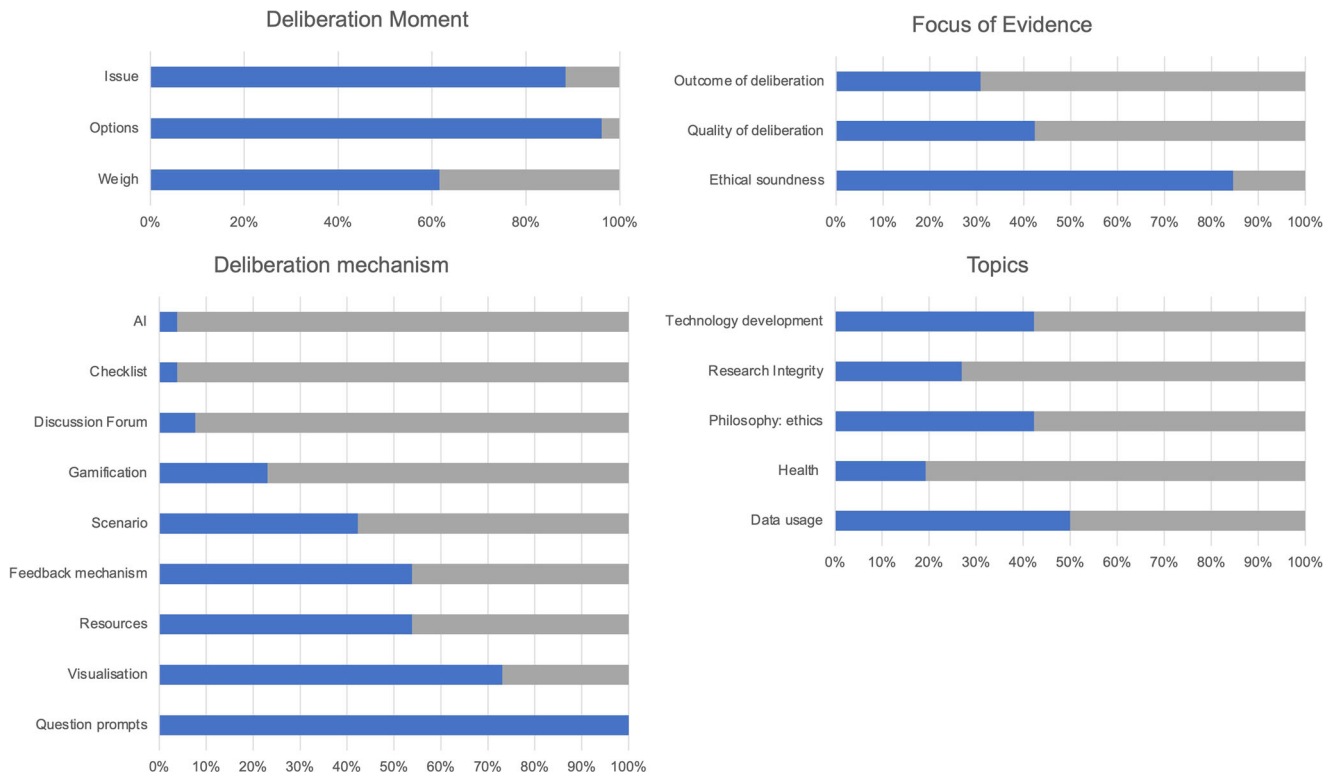
Fig. 3 Stacked bar charts showing the deliberation moment, deliberation mechanism, focus of evidence for effectiveness, and topics.

Notably, 11 of the 15 tools which had peer-reviewed publications were authored and developed by universities, highlighting how academic institutions are actively contributing to and shaping the landscape of ethical deliberation evaluation methodologies.

Discussion

In this study, we mapped 26 digital tools designed to assist users in identifying ethical issues, imagining potential resolutions, and weighing competing solutions. The findings revealed a rich tapestry of deliberation mechanisms, and validation methods. In the ensuing discussion, we now consider the broader implications of digitalisation on ethical deliberation. Specifically, we examine the potential gains and losses for both individuals and collectives who engage with these digital tools, all the while scrutinising how the mechanisms employed foster certain utilities and ideas. Echoing the insights of Wright and Street (2007), who explored the influence of the physical design of a parliamentary space on the nature of debate, we recognise that the design of digital tools profoundly shapes the type and quality of ethical deliberation.

To begin, consider the study's observation that conventional deliberation mechanisms, such as scenarios and checklists, persist in the digital realm. This finding implies that despite the advent of technology, conventional approaches to ethical deliberation maintain their significance. However, this outcome is not unexpected, considering that designers often incorporate familiar elements to facilitate the transition from paper-based practices to digital technologies (Legner et al., 2017). In this context, digital tools assume a complementary role, enriching and augmenting traditional methods rather than replacing them outright.

The study's mapping of digital tools unveils how digitisation allows for the integration of deliberation mechanisms and the addition of functionalities. Consider, for instance, the utilisation of scenarios, a mechanism long employed in narrative ethics because the contemplation of potential futures can spark ethical imagination and reflexivity (Baldwin, 2015). Digitising this mechanism allows for its combination with gamification elements and interactive visuals, thereby elevating appeal, and audience engagement. Similarly, checklists, historically instrumental in operationalising ethical guidance, are not without their challenges, as scholars caution against their potential limitations in stimulating critical reflection and reasoning (Madaio et al., 2020). Combining checklists with supplementary resources and personalised feedback mechanism in the digital realm can address this concern by fostering a more enriched and documented ethical deliberation process. Furthermore, through digital transformation, tools not only boast enhanced functionalities to improve the overall user experience but also have the potential for broader accessibility. In doing so, digital tools can potentially contribute to the well-being of communities or collectives by fostering shared values, understanding, and ethical considerations. Moreover, developers can leverage data on tool usage to inform design enhancements and contribute valuable insights to ethics research.

Digitalisation also enables the possibility of quantification. Here we refer to tools which provide users with quantified ethical assessment reports, appealing to ideals of neutrality, rigor, objectivity, and more credible decision making. Jasanoff (2005) refers to this as 'technologies of Hubris,' promising control in uncertain domains. However, scholars have raised concerns regarding the potential risks associated with quantification, including reductionism, bias, and persuasiveness (Saltelli and Di Fiore, 2020). To illustrate this, consider the tools in this study that used feedback mechanisms to display the percentage of users who concurred with a specific resolution, often without transparency regarding whose perspectives the numbers represent. Such feedback can side-line subjective, cultural, or contextual factors that are important for understanding different stakeholder perspectives. This trend towards reductionism and simplification mirrors a broader pattern in media technologies, whereby we see simplified and personalised content delivery on social media create

filter bubbles and echo chambers, contributing to societal polarisation (Pasca, 2023; Light et al., 2017). To address this, Jasanoff (2005) advocates for the development of 'technologies of humility,' encouraging reflection on ambiguity, indeterminacy, and complexity, moving beyond the use of a binary logic especially in the presence of reasonable disagreement.

The impact of technology on ethical deliberation goes beyond its capacity for quantification; it profoundly shapes the entire deliberative process, as evidenced by the dominance of tools designed for individual use. In our sample, most tools (73%, $n = 19$) were tailored for individual use, with only two incorporating discussion forums to facilitate collective deliberation. The inclination toward individualism can be attributed to the technology, as computers and mobile devices are primarily designed for single-user purposes. The limited utilisation of discussion forums could then stem from the need to monitor chats for preventing misuse and ensuring a platform conducive to constructive discourse. This underscores the inherent design and technical complexities associated with enabling collective online deliberation, demanding both a sufficient quantity and diversity of participants and effective moderation to uphold the quality of discourse (Wright and Street, 2007).

The observed variations in tools, encompassing diverse contexts, audiences, and intended uses, aligns with differences in the application of the dilemma or problematic approach. The dilemma approach, traced back to the Greek word 'lêmma' (meaning 'what one takes') and the prefix 'dís' (meaning 'two'), as Gracia (2003) suggests, revolves around presenting users with two opposing propositions. In our sample, The Trolley Dilemma game exemplified this, in that it asks users to choose between two contrasting options, highlighting a clash of values. Conversely, the problematic approach, derived from the Greek 'próblema,' signifies posing questions to be answered or solved, as described by Garcia (2001). Tools using this approach prioritised the process over the conclusion, emphasising the means rather than the end. For instance, the RRI Self Reflection Tool sought to address the entire process of ethical decision making and untangle the complexity of moral problems and solutions, rejecting the notion of a single, universal solution. These distinct approaches have different benefits and drawbacks, depending on the audience, setting, and goals. While the binary dilemma approach suits initial exploration, engaging a broad audience and encouraging them to contemplate their values, and decision making, it oversimplifies complex ethical matters and lacks the depth to capture real-world nuances that involve conflicting values. On the other hand, the problematic approach may appeal to users seeking an in-depth exploration of the intricate web of ethical considerations, values, and real-world contexts. However, it may also prove overwhelming for some users.

Lastly, let us reflect on the finding that one tool (EDEN AI) used ChatGPT, a Natural Language Processing (NLP), for facilitating ethical deliberation. This technology has recently gained significant attention in both public and academic discourse, drawing warnings from scholars about AI's inconsistent moral advice (Krügel et al., 2023) and ChatGPT's inaccuracy in bioethics question testing (Chen et al., 2023). Despite these concerns, the EDEN AI tool which we analysed in this study, takes a unique approach by presenting users with multiple ethical perspectives (e.g., virtue ethics or utilitarian ethics) to support problem solving and decision making. This allows users to weigh and compare outcomes, enhancing user autonomy by avoiding the imposition of a conclusive decision. However, it requires trust in the translation of underlying ethical theories and principles that inform the presented options. Much like many tools in our sample, the effectiveness of EDEN AI hinges on users trusting its ethical integrity, rooted in the translation of ethical perspectives (Laaksoharju, 2010).

This raises a broader question regarding how ethical deliberation tools must strike a balance between leveraging ethical theories

to guide deliberation, while refraining from exerting undue influence that might compromise user autonomy (Laaksoharju, 2010). As Light et al. (2017, p. 727) explain, the danger lies in developing tools that promote techno-paternalism, meaning "nudging users unthinkingly toward behaviour identified by others as positive, right or useful". Consequently, a critical evaluation of these tools' deliberative processes and their overall impact is necessary. For AI-driven tools, this underscores the importance of future research focused on evaluating their specific processes of deliberation and the consequences of their potentially widespread integration into other digital tools and technologies.

Limitations. The present study has several limitations that should be considered. First, the sampling method employed in this study may not have captured all existing tools in the field. This is because the landscape of technology tools is constantly evolving, and tools constantly emerge and become obsolete over time. Second, the sample consisted of only English language tools, limiting the generalisability of the findings to non-English speaking populations and contexts. Similarly, it is important to note that the sample size is small and the list of deliberation mechanisms not exhaustive. The taxonomy presented in this article should thus be considered primarily as a provisional framework to facilitate research on ethical deliberation tools, as our goal was to focus on providing a broad understanding of these mechanisms, rather than delve into the analysis of sub-types, the significance of between-mechanism variables, or specific design strategies employed to foster deliberation. Finally, as is typical with qualitative content analysis, there is the potential for bias in coding. To mitigate this issue and enhance the reliability of our findings, we employed best practice by involving two independent reviewers in the coding process.

Conclusion

This research offers a comprehensive exploration of the evolving landscape of digital tools for ethical deliberation. The analysis presented in this study serves as a valuable resource for bioethicists, researchers, educators, and stakeholders interested in ethical deliberation tools as it sheds light on how various mechanisms can facilitate structured deliberation, inclusive stakeholder engagement, and transparent documentation. Furthermore, the strategies employed by the analysed ethics tools can extend their applicability to diverse contexts, including policymaking, law, education, and business ethics. Nevertheless, a critical challenge remains for these tools: navigating the delicate balance between utilising ethical theories for guidance and preserving user autonomy. Further research is warranted to comprehensively assess the influence and efficacy of emerging technologies like ChatGPT across various domains, contributing to a broader understanding of their potential benefits and challenges.

Data availability

The datasets generated and analysed during this study are available in the supplementary files.

References

Awad E, Dsouza S, Kim R et al. (2018) The moral machine experiment. Nature 563(7729):59–64
Baldwin C (2015) Narrative ethics for narrative care. J Aging Stud 34:183–189

Beekman V, Brom FWA (2007) Ethical Tools to Support Systematic Public Deliberations about the Ethical Aspects of Agricultural Biotechnologies. J Agric Environ Ethics 20:3–12. https://doi.org/10.1007/s10806-006-9024-7

Blasimme A, Fadda M, Schneider M, Vayena E (2018) Data sharing for precision medicine: policy lessons and future directions. Health Aff 37(5):702–9

Bloom BS (1956) Taxonomy of educational objectives, handbook: the cognitive domain. David McKay, New York

Brown S, Mecklenburg A (2021) The next generation of data ethics tools. Open Data Institute. https://theodi.org/news-and-events/blog/the-next-generation-of-data-ethics-tools/. Accessed 22 Nov 2023

Chen J, Cadiente A, Kasselman LJ, Pilkington B (2023) Assessing the performance of chatgpt in bioethics: a large language model's moral compass in medicine. J Med Ethics

Diallo SY, Shults FL, Wildman WJ (2021) Minding morality: ethical artificial societies for public policy modeling. AI Soc 36(1):49–57

Garcia D (2001) Moral deliberation: the role of methodologies in clinical ethics. Med, Health Care Philos 4:223–232

Godin K, Stapleton J, Kirkpatrick S, Hanning R, Leatherdale ST (2015) Applying systematic review search methods to the grey literature: a case study examining guidelines for school-based breakfast programs in Canada. Syst Rev 4(1):1–10

Goodin RE, Niemeyer SJ (2003) When does deliberation begin? Internal reflection versus public discussion in deliberative democracy. Political Stud 51(4):627–649

Gracia D (2003) Ethical case deliberation and decision making. Med, Health Care Philos 6:227–233

Gutmann A, Thompson D (1997) Deliberating about bioethics. Hastings Cent Rep. 27(3):38–41

Hilliard LJ, Buckingham MH, Geldhof GJ et al. (2018) Perspective taking and decision-making in educational game play: A mixed-methods study. Appl Dev Sci 22(1):1–3

Hsieh H, Shannon S (2005) Three approaches to qualitative content analysis. Qualitative Health Res 15(9):1277–88

Ienca M, Vayena E (2021) Digital Nudging: Exploring the Ethical Boundaries. In: Véliz C (ed) The Oxford Handbook of Digital Ethics. Oxford Academic. https://doi.org/10.1093/oxfordhb/9780198857815.013.19

Ilten-Gee R, Hilliard LJ (2021) Moral reasoning in peer conversations during game-based learning: An exploratory study. J Moral Educ 50(2):140–165

James KL, Randall NP, Haddaway NR (2016) A methodology for systematic mapping in environmental sciences. Environ Evid 5:1–13

Jasanoff S (2005) Technologies of humility: Citizen participation in governing science. Springer

Jobin A, Ienca M, Vayena E (2019) The global landscape of AI ethics guidelines. Nat Mach Intell 1(9):389–399

Kaiser M, Millar K, Thorstensen E, Tomkins S (2007) Developing the ethical matrix as a decision support framework: GM fish asa case study. J Agric Environ Ethics 20(1):65–80

Kim TW, Werbach K (2016) More than just a game: ethical issues in gamification. Ethics Inf Technol 18:157–173. https://doi.org/10.1007/s10676-016-9401-5

Krathwohl DR (2002) A revision of bloom's taxonomy: an overview. Theory into practice. 41(4):212–218

Krügel S, Ostermaier A, Uhl M (2023) chatgpt's inconsistent moral advice influences users' judgment. Sci Rep. 13:4569

Kvalnes Ø (2019) The Navigation Wheel. Moral Reasoning at Work: Rethinking Ethics in Organizations: 49–60

Kwasnik BH (1999) The role of classification in knowledge representation and discovery

Laaksoharju M, Kavathatzopoulos I (2009) Computerized support for ethical analysis. In: Proceedings of the Eighth International Conference on Computer Ethics – Philosophical Enquiry. 425-437

Laaksoharju M (2010) Let us be philosophers!: Computerized support for ethical decision making. Dissertation, Uppsala University

Lawrence AM, Sherry MB (2021) How feedback from an online video game teaches argument writing for environmental action. J Lit Res 53(1):29–52

Legner C, Eymann T, Hess T et al. (2017) Digitalization: Opportunity and Challenge for the Business and Information Systems Engineering Community. Bus Inf Syst Eng 59:301–308

Light A, Shklovski S, Powell A (2017) Design for existential crisis. In: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems: 722–734

Lukianets N, Nekrutenko V, Pavaloiu A (2021) The Open Ethics Canvas v1.0.1 (1.0.1) Zenodo. https://openethics.ai/canvas/

Lyreskog DM, Pavarini G, Jacobs E, Bennett V, Mawdsley G, Singh I (2023) Testing design bioethics methods: Comparing a digital game with a vignette survey for neuroethics research with young people. AJOB Empir Bioeth 14(1):55–64

Madaio MA, Stark L, Vaughan JW, Wallach H (2020) Co-designing checklists to understand organizational challenges and opportunities around fairness in

AI. In: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. 1–14

Maner W (2002) Heuristic methods for computer ethics. Metaphilosophy 33(3):339–365

Marti P, Iacono I (2016) Anticipated, momentary, episodic, remembered: the many facets of User experience. In: 2016 Federated Conference on Computer Science and Information Systems (FedCSIS), Gdansk, Poland. 1647–1655

Moula P, Sandin P (2015) Evaluating ethical tools. Metaphilosophy 46(2):263–279

Mystakidis S, Christopoulos A, Pellas N (2022) A systematic mapping review of augmented reality applications to support STEM learning in higher education. Educ Inf Technol 27(2):1883–1927

Monge Roffarello A, De Russis L (2023) Nudging users or redesigning interfaces? Evaluating novel strategies for digital wellbeing through incontrol. In: Proceedings of the 2023 ACM Conference on Information Technology for Social Good. 100–109

Navarrete CD, mcdonald MM, Mott ML, Asher B (2012) Virtual morality: Emotion and action in a simulated three-dimensional "trolley problem". Emotion 12(2):364

Pan X, Slater M (2011) Confronting a moral dilemma in virtual reality: a pilot study. In: Proceedings of HCI 2011 The 25th BCS Conference on Human Computer Interaction 25. 46–51

Pasca S (2023) Enhancing Awareness of Filter Bubbles and Echo Chambers: The Role of Information Consumption Patterns and Political Orientation. Dissertation, University of Twente

Petersen K, Feldt R, Mujtaba S, Mattsson M (2008) Systematic mapping studies in software engineering. In: 12th International Conference on Evaluation and Assessment in Software Engineering (EASE)

Ramos J (2019) Eliza explores the dangers of on-demand digital therapy: The latest from Zachtronics is not a puzzle game, but a visual novel. Polygon https://www.polygon.com/reviews/2019/8/12/20798997/eliza-zachtronics-visual-novel-review-impressions. Accessed 22 Nov 2023

Saltelli A, Di Fiore M (2020) From sociology of quantification to ethics of quantification. Humanit Soc Sci Commun 7(1):1–8

Senghor AS, Racine E (2022) How to evaluate the quality of an ethical deliberation? A pragmatist proposal for evaluation criteria and collaborative research. Med, Health Care Philos 25(3):309–326

Sleigh J, Ormond K, Schneider M, Stern E, and Vayena E (2023) How interactive visualizations compare to ethical frameworks as stand-alone ethics learning tools for health researchers and professionals. AJOB Empirical Bioethics. https://doi.org/10.1080/23294515.2023.2201479

Stark A, Thompson NK, Marston G (2021) Public deliberation and policy design. Policy Des Pract 4(4):452–464

Ten Have H, Patrão Neves MD (2021) Deliberation. In: Dictionary of Global Bioethics, 393-394. Springer

Towne WB, Herbsleb JD (2012) Design considerations for online deliberation systems. J Inf Technol Polit 9(1):97–115

Travis R (2010) Bioshock in the cave: Ethical education in Plato and in video games. In Ethics and game design: Teaching values through play, 86-101. IGI Global

Wimalasooriya C, Licorish SA, da Costa DA, Macdonell SG (2022) A systematic mapping study addressing the reliability of mobile applications: The need to move beyond testing reliability. J Syst Softw 186:111166

Wong RY, Madaio MA, Merrill N (2023) Seeing like a toolkit: how toolkits envision the work of AI ethics. Proc ACM Hum-Comput Interact 7(CSCW1):1–27

Wright S, Street J (2007) Democracy, deliberation and design: the case of online discussion forums. N. media Soc 9(5):849–869

Xenos M, Velli V (2018) A serious game for introducing software engineering ethics to university students. In: International Conference on Interactive Collaborative Learning. 579–588

Zhang-Kennedy L, Chiasson S (2021) A systematic review of multimedia tools for cybersecurity awareness and education. ACM Comput Surv (CSUR) 54(1):1–39

Zhuo TY, Huang Y, Chen C, Xing Z (2023) Exploring ai ethics of ChatGPT: A diagnostic analysis. Arxiv preprint arxiv:2301.12867

Zohud T, Zein S (2019) A systematic mapping study of cross-platform mobile apps. J Comput Sci 15(4):519–536

Zuber N, Gogoll J, Kacianka S, Pretschner A, Nida-Rümelin J (2022) Empowered and embedded: ethics and agile processes. Humanit Soc Sci Commun 9(1):1–13

Author contributions

Competing interests

The authors declare no competing interests.

Ethical approval

Ethical approval was not required as the study did not involve human participants.

Informed consent

Informed consent was not required as this study did not involve human participants.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1057/s41599-024-02629-x.

Correspondence and requests for materials should be addressed to Joanna Sleigh.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.