Defining socially disruptive technologies and reframing the ethical challenges they pose

Joel Anderson [a,*], Jeroen Hopster [a], Björn Lundgren [b]

[a] *Ethics Institute, Utrecht University, Utrecht, the Netherlands*
[b] *Centre for Philosophy and AI Research, Friedrich-Alexander-Universität Erlangen–Nürnberg, Erlangen, Germany*

ARTICLE INFO

ABSTRACT

Socially Disruptive Technologies (SDTs) loom large in public debate, yet scholarly discourse on the ethical implications of social disruption is still in its infancy. This article makes two contributions to advance this discourse. First, we propose and defend a new definition of SDTs that allows for classification of those technologies that warrant further ethical analysis, specifically in virtue of their socially disruptive nature (among the examples we discuss are deepfakes, cultured meat, birth control technologies). This is the applied value of the framework we offer: to offer guidance in identifying which technologies require specific scrutiny as SDT, and guidance in identifying an initial set of ethical tools to accompany such analysis. Second, we reframe the ethics of social disruption by highlighting how SDTs pose challenges to capacities for normative orientation and joint action-coordination and by foregrounding the potentially stratified availability of the resources needed to overcome or mitigate these challenges. We argue that although the burdensome disruptiveness means that SDTs characteristically have, at least in a "narrow" sense, a *pro tanto* negative valence, they may nonetheless turn out to contribute to significant social and moral progress. The ethical concerns raised by SDTs require an approach that is sensitive both to the challenges inherent in the disruption and to its eventual outcome.

1. Introduction

The term "disruptive technology" was introduced by Christensen (1997) as part of his theory of business innovation, but in recent years it has been further developed in connection with a variety of topics, such as "social disruption" (Schuelke-Leech, 2018; Hopster, 2021a; O'Neill, 2022; Choi & Moon, 2023; Van Grunsven & Marin, 2024; Lane, 2025), "moral disruption" (Baker, 2013; De Proost & Segers, 2023; Hermann, 2025; Nickel, 2020; Rueda et al., 2022), "legal disruption" (Brownsword, 2019; Liu et al., 2020; Maas, 2022), and "conceptual disruption" (Hopster & Löhr, 2024; Löhr, 2023; Marchiori & Scharp, 2024). Diverging from disruptive innovation theory's focus on economic markets, this new wave of scholarship examines how emerging technologies challenge social institutions, legal frameworks, conceptual categories, moral norms, and beliefs about reality.

As is often the case with a fast-emerging field of discourse, the lack of consensus on a definition can leave people talking past each other. Some authors have identified characteristic features of "socially disruptive technologies," hereafter, "SDTs" (esp. Hopster, 2021a; see also Bovenkerk & Boersma, 2023; Lane, 2025, Hermann, 2025). But a strict definition of SDTs, in the classical philosophical sense, is still lacking (e. g., Hopster, 2021a focuses on prototypical characteristics of SDTs, rather than necessary conditions, and there is no claim of joint sufficiency). Some scholars have even questioned whether socially and legally disruptive technologies can be defined *at all* (e.g., Carlsen et al., 2010; Liu et al., 2020), while others have criticized the "hype" around disruption (Daub, 2020).

We hold that sustained reflection on the nature of social disruption and associated technologies is an important endeavor, and that developing an adequate definition is likely to be conducive to appropriately articulating and assessing the ethical significance of SDTs. In this article, we offer a new definition of SDTs to support this claim and to make progress in connecting the discourse on social disruption. Our definition can help to pick out technologies worthy of further ethical scrutiny. That is, we take as one of the aims of the definition to identify a subset of technologies that warrants special attention in virtue of its disruptive features. This is not to say that SDTs should always be resisted as leading to bad overall outcomes (an assumption we oppose in section 3.4.2). Nor do we follow the techno-optimistic hype and uncritically praise the aspirations of disruptors. Instead, we argue that SDTs should be assessed in

---

a context-sensitive manner, with a distinct focus on the challenges to the capacity for normative orientation and joint action-coordination they pose. Methodologically, we are engaged here in conceptual analysis, focusing on formulating necessary and sufficient conditions for SDTs through careful examination of shared features in paradigm cases (taken from both current and past cases). This approach differs from empirical social science methodologies that might seek measurable criteria or operational frameworks. Instead, we aim to provide conceptual clarity that can inform both theoretical discussion and practical application, recognizing that different methodological approaches offer complementary insights into technological disruption.

Our approach in working towards a definition is to first propose and defend key desiderata for a definition of SDTs (section 2), before elaborating on our proposal to understand SDTs in terms of their challenges to capacities for normative orientation and joint action-coordination (section 3). Next, we highlight some distinctive ethical concerns raised by SDTs and counter the objection that our account misconstrues these concerns (section 4). Section 5 concludes.

2. Defining socially disruptive technologies

To develop our definition of SDTs, we begin by identifying key desiderata that any adequate definition must satisfy and then provide a formal definition that captures the essential features of SDTs while remaining applicable across diverse contexts (in 2.2). In 2.3, we then examine paradigm cases of technologies commonly recognized as socially disruptive and discuss how they fit our definition.

2.1. Desiderata of a definition of socially disruptive technologies

Numerous technologies have been labeled "socially disruptive," spanning from emerging technologies (e.g., quantum computing, blockchain, and gene drives) and recent innovations (e.g., internet-based social platforms, synthetic fertilizers, and oral contraceptives) to technologies from the distant past (e.g., plows, astrolabes, and gunpowder) (see, e.g., Schuelke-Leech, 2018; Giovanola, 2023; Hopster et al., 2022). Yet, while their classification as SDTs is often intuitive, the literature lacks an account of what the necessary and jointly sufficient criteria are for a technology being not simply innovative but disruptive. Hopster's (2021a) account comes closest: it outlines seven criteria in virtue of which a technology could be said to be more rather than less disruptive, but it does not offer an explicit definition of SDTs. The present article complements his account by offering a definition of SDTs, consonant with these previously specified criteria for disruptiveness.

First, some expectation management. We take to heart Aristotle's dictum that clarity should be sought to the extent that the topic admits of it. SDTs undoubtedly comprise a somewhat undifferentiated set, forming not a natural kind but a social kind (Mason, 2016). Nevertheless, a definition of SDTs can help to synthesize the discourse and provide a common reference point that connects recent discussions of social disruption. Similar definitional attempts have helped to clarify the concept of "emerging technologies" (Rotolo et al., 2015) and "digital technologies" (Baier et al., 2023). Furthermore, offering a productive definition may involve stipulative elements that can catalyze theory-building. Our proposal should be understood accordingly: it builds on existing discourse, while foregrounding a distinctive angle on social disruption that we regard as fruitful for ethics of technology and normative political theory.

To set aside another potentially controversial issue at the outset, our definition does not commit to a detailed proposal of what qualifies as a "technology" but relies on an informal understanding of it. There are different levels at which technologies can be individuated (Brey, 2017) and different proposals as to how their nature should be understood. For instance, *materialist* theories think of technologies as essentially being material artifacts, whereas *institutionalist* theories also include social institutions in their scope (Danaher & Sætra, 2023). We do not think that

we need to take a specific stance on this or related distinctions (Schatzberg, 2018), since our definition explicitly presents the broader technosocial context as being co-constitutive of a disruption and, hence, of an SDT. Furthermore, the hallmark of an SDT, on our view, does not stem from features intrinsic to the technology, but rather from its social implications.

Finally, a good definition of SDTs needs to strike a balance. It should be sufficiently capacious to encompass the enormous variety of what can pass as a technology and as a social disruption. At the same time, not just any change introduced with technologies should count as "socially disruptive," if the phrase is to be useful. To delineate scholarship on SDTs as a distinctive and normatively interesting field, it is particularly important to highlight the substantial ways in which technologies have the potential to alter and destabilize interpersonal, institutional, psychological, cultural, and conceptual circumstances in ways that have profound implications for the wellbeing of humans and the functioning of social institutions. Central to our account, then, is the idea that social disruption creates challenges that make it difficult to go on as before, forcing agents to adapt, cope, or resist. Foregrounding the agency of those facing disruption sharpens the focus on the complex of ethical concerns engendered by SDTs. Building on these properties, our definition can help to pinpoint technologies that warrant further ethical analysis, in virtue of their socially disruptive nature.

2.2. Definition

We propose the following definition:

Socially Disruptive Technology (SDT): X is an SDT, in a given social context C, if and only if X is a technology that, if developed or implemented, generates or significantly amplifies substantial challenges to the level of normative orientation and joint action-coordination available in C.

There are several aspects worth highlighting of how we propose to construe SDTs with this definition, which we discuss in section 3: contextual specificity (3.1); the destabilization of capacities for normative orientation and joint action-coordination (3.2); the positive or negative valence of disruption (3.3); the tractability of challenges and potential for coping effectively with them (3.4); the "substantial" character of challenges (3.5); and the often stratified availability of coping resources (3.6). Note that this definition aims to provide a conceptual framework rather than an empirical measurement tool. Rather than providing itself the operational metrics for quantifying disruption, it establishes the conceptual foundations necessary for such empirical work. We view the contextual nature of our definition not as a limitation but as reflecting the genuine variability in how technologies affect different societies.

2.3. Paradigm cases of SDTs

Before elaborating on these points, consider three examples which illustrate the distinctive features of SDTs foregrounded by our definition.

2.3.1. Deepfakes and generative AI video imagery

A first example are deepfakes and technologies that can be used to fabricate audiovisual material. Before convincing forms of such technologies became widely available, audiovisual material fulfilled a crucial function in providing reliable and trusted evidence of events. From police body camera imagery in criminal courts to videotaped interviews with politicians, having events "on tape" regularly served as conclusive evidence about what really happened. The possibility that video imagery is fabricated creates a substantial challenge to the crucial societal task of epistemic coordination of a shared reality (Rini, 2020). At present, there is a good deal of discussion about whether and how this challenge can be surmounted to restore reliable means of authentication. But for now, the social and institutional practices for verifying

information are disrupted. This disruption leaves people disoriented across a wide range of contexts: not only does it raise the epistemic burden of differentiating between what is real and fake, but it also challenges once-held certainties regarding evidential norms (Fallis, 2021; cf. Habgood-Coote, 2023 for an argument that deepfakes are not that concerning; see also Öhman, 2024 for a discussion of how knowledge production is changing with generative AI). In a parallel development, the use of these technologies to generate new video imagery is destabilizing norms around creativity (as stunning videos can be created without the skills or effort previously required) and exploitation (as GenAI videos are fabricated without any payment to actors or copyright holders whose work has been used in training those models). As the debates rage on, many now find themselves disoriented on questions of which uses of these technologies should count as exploitative or genuinely creative. Lastly, deepfakes and generative AI raise a plethora of issues that we cannot deal with in full, some of which can be disruptive, such as the idea that fake content in deepfakes can violate our right to privacy (Lundgren, 2025).

2.3.2. Cultured (Lab-Grown) meat

The emergence of cultured meat fits the classic model of a "disruptive innovation" that shakes up established ways of doing business (Christensen, 1997), but it also represents a deeper challenge to fundamental aspects of agriculture, food consumption, and animal ethics. Animal husbandry has been a cornerstone of human civilization for millennia, deeply embedded in cultural practices, economic systems, and personal identities. The possibility of producing meat without animal slaughter challenges many of the justifications given for conventional meat consumption (Hopster et al., 2022). At the same time, for vegans and vegetarians, it is a disorienting question whether eating cultured meat violates their principles. The prospect of a future in which cultured meat is produced affordably and at scale raises profound challenges to global food systems, potentially redefining agricultural practices, supply chains, and regulatory frameworks, particularly in light of uncertainty about consumer acceptance, environmental impact, and long-term health effects. It may well be that challenges engendered by this technology will, once the challenges are met, lead to outcomes that are preferable to current arrangements, but even so, overturning current practices may involve setbacks for various stakeholders, aggravate social friction, and raise uncertainty about the demands of a just transition. See, for example, Siegrist et al., 2024; Mendly-Zambo, Powell, & Newman, 2021 on some of the complexities of transitioning to cellular agriculture.

2.3.3. Birth control technologies

The introduction of oral contraceptives and IUDs in the 1960s is a paradigm case of an SDT. Prior to their availability, reproduction was still largely regarded as a largely unavoidable consequence of sexual activity, shaping societal expectations around marriage, family planning, and women's roles. The pill's capacity to reliably separate sex from procreation disrupted deeply ingrained social and religious norms and empowered women. Individuals and institutions struggled to reconcile traditional values with the new possibilities for sexual autonomy and family planning. In many low-income countries, this disruption intersected with concerns about population growth, economic development, and cultural preservation, leading to complex debates about the role of family planning in national progress. The technology challenged traditional power structures and gender dynamics, particularly in societies where women's status was closely tied to fertility. The redefinition of women's social and economic roles, enabled by increased control over fertility, necessitated widespread reevaluation of gender norms, labor practices, and development strategies (Benagiano et al., 2007; Whicker & Kronenfeld, 1986). At the same time, unequal access to contraceptives between and within countries has added another layer of normative uncertainty and contestation. This is reflected, for example, in the opposition from women of color in the US, many of whom saw birth control

technologies as part of the legacy of white control over their sexuality (Nickel et al., 2022). To take things a step further, the disorienting effects of reproductive technologies may soon extend to ectogestative (artificial womb) technology, further challenging concepts of birth, parenthood, and human development – fundamental categories that structure both normative orientation and social practices (Frank et al., 2023).

Having established our definition and illustrated it through examples, we will now examine six key aspects that clarify how this definition operates in practice and what it reveals about the nature of SDTs.

3. Aspects of the definition

The above examples exhibit several characteristic features of SDTs: they involve uncertainty and contestation, individual and collective disorientation, challenges that stem from the differential ability to cope with disruptions across groups and societies or to transform the disrupted circumstances into durable social progress. Having provided these initial illustrations, in this section we further clarify the different terms of our definition and argue for their merits.

3.1. The contextual specificity of disruptiveness

Our definition emphasizes the contextual character of an SDT. In part, this is a familiar point about the sense in which every technology is a techno-socio-moral assemblage, which takes its shape in human practices, against a material and institutional background, and which is infused with symbolic interpretations (cf. Feenberg, 2017; Jaeggi, 2018). Any assertion about the disruptiveness of a technology is underspecified without reference to the circumstances in which it is deployed. This point is crucial for appreciating the historicity of disruptiveness: what was disruptive at one time may not be so at another. While the emergence of the railway as a means of transportation was hugely disruptive with respect to a range of features of mid-19th century life and society across the world, "train disruptions" now refer to an entirely different and all-too-familiar aspect of modern life.

Disruptiveness can also vary with cultural or societal contexts: some technologies may be particularly disruptive in the highly digitalized Global North, others in the mineral-rich Global South. We can speak of disruptions with a more local or global scope, as well as disruptions that pertain to specific groups or communities: while the emergence of cochlear implants was not disruptive for human populations at large, it has been profoundly disruptive within the deaf community (Scully, 2019). Even more local contexts might be thought of, though disruptions that merely pertain to localized interpersonal interactions and do not extend to "social contexts" more generally fall outside our scope.

In speaking of circumstances ("C"), the relevant criterion for assessing disruptiveness is that of those who are subject to disruption, not necessarily where the technology is developed or implemented (e.g., nuclear weapons being implemented and developed in one country, can lead to social disruption elsewhere). For example, in a 22-country comparative study, Fritz et al. (2024) found that geoengineering technologies create fundamentally different disruptions across political cultures, with publics conceptualizing their role in governance ranging from passive recipients to active decision-makers. Note that, since the ethical challenges of SDTs are not necessarily restricted to the local context of implementation, this raises further concerns for implementation and how to apply ethical recommendation for guidelines, standards, and regulations (cf. Lundgren, 2021, 2023a; 2023b). This is one of the respects in which SDTs show their 'second-order' disruptive nature (Hopster & Maas, 2023): they challenge the standing of the existing ethical, legal, and social norms and institutions to which we would otherwise appeal in addressing these disruptions.

Just as the disruptive nature of an SDT can vary over various contexts, so can the specific ethical concerns, because of variations in why

these disruptions matter to the people affected. For example, the emergence of artificial womb technologies is likely to raise different ethical concerns (and with different intensity) in cultures with strong, often religious views on abortion. The respective attachment to pre-existent cultural practices and the weight assigned to the associated norms are key contextual variables influencing the degree of moral disorientation to which SDTs give rise. Disruptiveness and its ethical evaluation also depend on the unequal and intersectional demographics of SDTs' impact, which has also been recognized in prior work on SDTs, and has been discussed under the header of 'differential disruption' (Nickel et al., 2022). Technosocial disruption affects different groups in different ways, which crucially depends on the availability of "coping resources" (we expand on this in section 3.4). Whether one belongs to a well-established elite or to a marginalized population may influence whether a new technology is abhorred as threatening privilege or welcomed as enhancing agency.

All this variability serves to highlight, once more, that there is nothing *intrinsic* about a given technology that makes it socially disruptive. As with other relational concepts such as "affordable," "vulnerable," or "nutritious," identifying what makes a technology "disruptive" requires attending to the salient aspects of the context, including the degree of entrenched reliance on previous technologies and associated norms and the availability of resources for adapting to or resisting the disruption.

3.2. The social character of what SDTs disrupt

A social disruption may generate positive or negative outcomes, or both. It may jeopardize or strengthen human rights, well-being, and our living conditions. While we recognize that these are important aspects of social disruptions, our definition has a different target phenomenon: it foregrounds the distinctively *social* character of disruption, which challenges agents' normative agency and may affect the fluency and efficacy of cooperative social structures.

To capture the distinctively social character of the disruption, our definition takes an explicitly *functionalist* turn that focuses on the close connection between "disruption" and "breakdown" that is often prominent in dictionary definitions (Hopster & Löhr, 2024). This component of "breakdown" helps to distinguish social disruption from mere social change. A disruption is a point at which a given social arrangement or practice no longer functions as before (for a related but distinct – and distinctively Wittgensteinian – take on "not being able to go on the same way", see Hermann, 2025 as well as Lane, 2025). The practice no longer serves its initial purpose (which can include unjust or malignant purposes), at least not without further investigation and possible modification. There is no going back to the prior state – disruptions are resistant to reversibility – but neither is there a clear path forward. Recall our earlier example of relying on audiovisual material as a source of uncontested evidence: the existing practice is disrupted, its function breaks down, leaving its practitioners – at least initially – perplexed about how to proceed. Note that we are not presupposing that there is only one solution: multiple "regimes" of sense-making and action-coordination may provide feasible solutions within a given social problem-space (Anderson, 2014).

To capture this "functional breakdown," we propose to characterize the specifically social character of SDTs' disruptiveness in terms of its effects on the ability of agents affected by them to meet key functional demands in social contexts, either in jointly coordinating action or orienting themselves in social-normative space. These are two characteristic and deeply intertwined features of a well-ordered society which SDTs challenge: joint action-coordination and normative orientation, both of which will be explained below.

3.2.1. Challenges to joint action-coordination

As much recent work on human evolution highlights, taking advantage of opportunities for cooperation and social learning has been key to the success of our species (e.g. Heath, 2006; Henrich, 2016; Sauer, 2024), including capacities for conflict-resolution. We propose that SDTs earn their designation precisely by disrupting these foundational capacities for collective agency, undermining the social practices, institutions, and norms that allow societies to meet functional demands for collective problem-solving and coordinated action. Paradigm cases are those in which problem-solving, action-coordination capacities built up in institutions and social practices are upended by technological developments. To extend the deepfakes illustration discussed in section 2.2.1, think of how social media technologies have upended the way in which fact-checking and the verification of political news has been decentralized away from the control of established journalists at leading newspapers or networks (Otto and Köhler, 2018). Or consider the dynamics of political polarization which have been exacerbated, among others, by the same social media (Hopster, 2021b), yielding challenges to core democratic institutions, even to the point that election results are no longer trusted and accepted.

There are a variety of theoretical frameworks that could be used to elaborate various dimensions of action-coordination that SDTs characteristically disrupt, something we can only sketch here. One theoretical source can be found in work on collective intentionality and agency. Margaret Gilbert (2014) has shown how societies depend on networks of joint commitments that enable groups to function as plural subjects with genuine collective agency. These commitments create a normative infrastructure – shared understandings of who is obligated to do what – that makes coordinated action possible.

Similarly (and in ways that anticipate section 3.2.2 below), "neo-institutionalist" scholars such as W. Richard Scott (2014) have shown how coordination depends on several categories of institutional supports: regulative (rules and enforcement), normative (values and expectations), and cultural-cognitive (shared understandings and categories). SDTs may destabilize all three simultaneously. When deepfake technology undermines audiovisual evidence, it doesn't merely create a technical problem—it erodes the regulative frameworks of legal proceedings, the normative expectations of journalistic integrity, and the cognitive categories through which we distinguish truth from falsehood. This multi-dimensional disruption explains why societies struggle to exercise their capacity for collective agency – the ability to form joint commitments that enable collective action.

Further, as Elinor Ostrom's empirical work (Ostrom, 2010) demonstrates, effective societies develop overlapping, polycentric systems with specific design features: clearly defined boundaries, collective-choice arrangements, graduated sanctions, and conflict-resolution mechanisms. The need for legal guardrails and effective institutional guidance is a recurring theme in discussions of technology. These governance arrangements often emerge slowly and in ways that are responsive to specific contexts. Once they are undermined, it becomes a major challenge to restore effective collective agency, particularly given the novelty of SDTs and the "pacing problem" they generate (Marchant, 2011). Social media platforms exemplify this: they've created new spaces for interaction faster than societies can develop appropriate governance mechanisms, leading to regulatory vacuums around data privacy, content moderation, and algorithmic accountability. The absence of clear boundaries (who is a member?), collective-choice arrangements (who makes rules?), and conflict-resolution mechanisms (how are disputes resolved?) creates coordination breakdowns from online harassment to election disinformation. And yet another example of the disruption of action-coordination can be found in AI-accelerated innovations in synthetic biology, where automated design-build-test-learn cycles bypass traditional scientific deliberation, creating what Groff-Vindman et al. (2025) call an 'interpretability crisis' where biological designs are generated without transparent reasoning, undermining established norms of scientific accountability.

A recurrent theme in these theoretical discussions is the importance of the alignment of *mutual expectations* for the successful coordination of joint action. Such alignment may be challenged by SDTs, whose rapid

and unpredictable development can raise very different expectations among social stakeholders (think of the polarized discourses on AI singularity). SDTs may also be interpreted and conceptualized in different and sometimes contested terms, fueling ambiguity towards existing practices and their moral status (consider the question whether or not sexual intercourse with a sex robot can qualify as adultery). Another aspect of joint action-coordination is the existence of *recognizable roles* performed by core institutions and agents in society. SDTs frequently contribute to the erosion and even the obsolescence of existing institutions, as well as a shift in social roles (think of 'users' of GenAI who simultaneously become 'workers' by providing data to the industry; or think of parents who can monitor the school performance of their children in real-time, which encourages them to take on responsibilities previously held by teachers).

In the terminology of Boltanski and Thévenot (2006), action coordination depends on the existence of shared "orders of worth" – legitimate principles of evaluation in a variety of domains (civic, market, industrial, domestic, inspired, fame) that provide normative grammar for justifying actions and resolving disputes. Among others, there is a need for legitimate authorities (formal or informal) with the requisite "normative powers" – the collective capacity to assign statuses and making determination of what meets various standards. For example, the economics of the art world require reliable ways of assigning authenticity and authorship to works of art. Against the background of this need, SDTs often precipitate "critical moments," in which established orders of worth no longer furnish adequate resources for coordination. For example, the emergence of non-fungible tokens (NFCs) has introduced new ways of assigning the normative status of "being the author of a work of art," and upending established authorities for assigning the status of being an authentic owner.

3.2.2. Challenges to the capacity to orient oneself in socio-normative space

The second dimension of disruptive social impact that our definition highlights is *normative disorientation*. The confrontation with new technologies often destabilizes established normative categories, principles, and concepts with which members of society are able to orient their decision-making and their interpretation of their experiences (see, e.g., van de Poel et al., 2023; Congdon, 2024). New technology can scramble our moral radar and undermine our confidence in our moral judgments. This can occur at an individual level, with some people facing greater difficulties finding their way in the wake of transformed vocabularies or evaluative schemas, but SDTs can also affect a jointly realized capacity for orientation on the basis of shared points of reference. In this sense, *conceptual* disruption is also an important factor that can contribute to what we term "normative disorientation" (Hopster et al., 2023; Löhr, 2023; Marchiori & Scharp, 2024).

On our usage, the term "normative orientation" refers not primarily to *agreement in basic beliefs or judgments* (though this can be relevant; see also Lane's 2025 discussion of deep agreement), but rather to *the way established and interlinked norms provide guidance and enable fluent action in a given social context* (Henrich, 2016; Ostrom, 2000). This also illustrates the interconnections between the role of norms both orientation (discussed here) and in action coordination (discussed in 3.2.1). Formal and informal institutions, codified rules, and shared concepts underpin and stabilize these normative orientations, as do material affordances mediated by technology. These landmarks guide our paths, even when we choose to head in very different directions. In addition, normative orientation pertains to the autonomous agency of individuals: their capacity to be the author of their central evaluative commitments and their having a real capacity to act upon them.

Normative orientations are omnipresent but often implicit: norms may be mutually acknowledged and even internalized, without the need to explicitly reflect on their contents (Swierstra, 2013). This changes when normative orientations are disrupted, and the fluent and skillful norm-following characteristic of ordinary practice comes to a halt. Reflection may be needed as to how to go on, when one's prior ability to

solicit normative affordances is undermined or when new normative affordances emerge. A move towards normative scrutiny is a hallmark of social and conceptual disruptions induced by technology: they upset established normative orientations and provoke explicit reflection on how to go on (Haslanger, 2020).

While the challenges provoked by SDTs typically involve breakdown of existing normative orientations, there are also cases where SDTs give rise to an entirely new normative orientation, as a new domain of human interaction emerges which previously did not exist (Harbin, 2016). Possati (2023) argues, for example, that quantum computing requires entirely new conceptual tools for ethics, as existing frameworks around data transparency and algorithmic accountability become meaningless when applied to quantum systems. Or consider the newly emerging norms of online social media, which have filled a normative lacuna: they pertain to a new realm of human action and interaction that did not exist two decades ago. Or consider how contraception has contributed to new norms around sex. The process of acquiring shared normative orientation in the online realm has certainly been challenging, as suggested, for instance, by the various harms associated with online 'moral fog' (Cocking & van den Hoven, 2018). The challenge, in this case, does not so much pertain to an existing normative orientation that has been overturned, but rather to generating newfound orientation in the online domain where previously there was none – although a newfound capacity for orientation in one domain, might well be associated with a receding capacity to do so elsewhere.

3.3. The actual and potential challenges associated with SDTs

We have characterized disruptions in terms of breakdowns. On our definition, however, it is not *actual* breakdowns, but the *potential* for breakdown that suffices for classifying technologies as SDTs. It can be apt to label technologies "disruptive," owing to their disruptive potential, even when they are *never* actually implemented and have no impact in *any* context. This accommodates the fact that certain ethical concerns are raised already by the potential impact of SDTs that are still in early stages of development – think of solar geo-engineering technologies, quantum computing, nuclear fission, blockchain fintech, the metaverse, or space mining. A good definition should enable one to identify future SDTs – that is, technologies that are anticipated to be highly disruptive in a future context, even if it is uncertain whether this future possible context will obtain.

For example, consider the possibility that governments successfully maintain a ban on high-tech solar radiation management technologies (as proposed, for instance, by Biermann et al., 2021). The rationale for this ban derives, in large part, from the socially disruptive nature of the technology: proponents of the ban argue that our global system of government is ill-suited to direct solar geo-engineering, and that the technology would have a detrimental impact, if it were to be implemented. Now, suppose that the ban is successful, such that solar geo-engineering technologies are never implemented or even fully developed, and the disruptive impact is prevented. Clearly, we should still think of solar geo-engineering as an SDT: its socially disruptive *potential* was the very reason for blocking its development.

3.4. What makes a challenge substantial?

In some sense, the introduction of any technology comes with some challenges. To avoid having the class of SDTs become meaninglessly broad, our definition requires that the challenges introduced by a technology be "substantial." There are several points to be made about this.

3.4.1. Criteria for substantial disruptiveness

Challenges are substantial when a technology impinges on normative orientations or joint action-coordination in a thoroughgoing way, which, for example, can be spelled out in terms of the "disruptiveness

criteria" identified by Hopster (2021a), including such factors as "reversibility of impacts," the "pace of change," or the "extent of uncertainty." As such, the 'substantiality' threshold is vague. However, the lack of sharp borders does not mean that there aren't clear cases. Our suggestion is that a disruptive challenge is substantial in cases where a solution isn't readily or easily available; communities in social context C are not in a position to adequately respond to the challenge. This implies that the resources for responding available to a community are crucial in determining the disruptiveness of SDTs, a point we discuss in more details in section 3.4.3.

Our primary reason for adding a substantiality threshold to our definition is to justify the exclusion of examples that intuitively do not carry sufficient weight to qualify as SDTs. Consider market disruptions that have little social significance, even if they involve technological changes that are disruptive to limited branches of business and industry. For instance, the introduction of anti-lock brakes in automobile technology during the 1970s (Carlsen et al., 2010) disrupted the car manufacturing industry, affected many people, and may well have led to some individual confusion. But it did not challenge normative orientations or action coordination in any significant way. Similarly, changes to the diameter of computer disk drives during the 1980s constituted a major disruption to the hard-disk drive industry (Bower & Christensen, 1995), but these did not substantially affect social practices, let alone challenge normative orientations.

What makes a disruption substantial is not merely the extent of the impact or the number of people affected by it. Here we diverge from Schuelke-Leech's (2018) understanding of "second-order disruption," according to which SDTs are defined in terms of the magnitude of a technology's actual impact, whose consequences "ripple through society" (idem, p. 261). In our view, it would be overly restrictive to make the extent of impact either a necessary or sufficient criterion of disruptiveness, although we grant that the range of impacts can be contributing factor (see Hopster, 2021a).

Finally, we also take it that an SDT can contribute to an already existing substantial challenge without being the original cause of it. If the contribution is significant enough, then the technology can count as an SDT. Future empirical research would be needed, however, to develop context-specific indicators for substantiality, building on the conceptual foundation we provide here.

3.4.2. The narrowly negative but potentially beneficial character of disruptions

In using the phrase "substantial challenges" we are distancing ourselves from the claim that SDTs are defined by their negative impact. Indeed, one of the central difficulties in finding an adequate definition of SDTs lies in acknowledging the hugely beneficial effects that some technological disruptions can ultimately have, without losing sight of the ways in which disruptions have something undesirable about them. We believe both aspects must be accommodated by a definition of SDTs: even when we know disruptions to have significant costs, burdens, and inconveniences, we may nonetheless view the costs to be worth paying, for even burdensome and taxing disruptions can lead to improvements that could not have been achieved otherwise. Indeed, disorientation may be an essential part of gaining moral insight (Harbin, 2016) or dispelling oppressive ideologies (Hayward, 2020), and very often reaching a better cooperative equilibrium requires a disruption of the prior, suboptimal equilibrium (Skyrms, 2010).

To square the circle, we make two moves. First, we propose to read our definition in light of a distinction between a narrow and a wide construal of how beneficial or harmful a disruption is. Narrowly construed – that is, looking at a narrow set of concerns raised by the process of disruption and its immediate implications – social disruptions are typically negative: they are costly, disorienting, and burdensome. As "narrowly negative" phenomena, SDTs provoke destabilizing changes that force a search for new ways to respond, leaving a temporary vacuum with regard to both normative orientation and action-coordination.

The social disruption SDTs threaten to bring about involve no longer knowing how to go on, not yet seeing a feasible solution, or being unwilling to make the requisite changes. Narrowly construed, periods of social disruption are periods of uncertainty, given how difficult it is to predict how things will turn out.

It is important to emphasize that people's attitudes towards this distinctive aspect of disruptiveness can vary widely. For example, techno-optimists or risk-tolerant entrepreneurs may relish the uncertainty and disorientation, even when it is unclear how things will play out. And for those who have been marginalized and oppressed by the status quo, the fact that key societal functions have been overthrown may be experienced as liberating. Our claim is not that these reactions are irrational or that stability necessarily deserves priority. But even for those who consider the price to be worth paying, the costs are often real. In that sense, even in situations where the leap into the unknown (a period of burdensome disruption and disorientation) is enthusiastically embraced, it can still be aptly understood to be burdensome, risky, and taxing, where it is unclear what the way forward is and whether it will lead to improvement or to disaster.

On a broad construal, by contrast, disruptions are assessed on the basis of how things eventually turn out, for better or worse. SDTs may end up being hugely beneficial, leading to less oppressive or more optimal ways of securing normative orientation or joint action-coordination. This is not a given, of course: progress is possible but not inevitable. Indeed, the eventual implications of SDTs are typically a mixed bag, involving advances with regard to some aspects of human wellbeing and social flourishing, and setbacks with regard to others. This is also recognized in historical and political treatises which take a birds-eye view on disruption (e.g. Fukuyama, 2017; Zakaria, 2024): while in the long run progress often does ensue, this comes with certain losses, which cannot always be undone.

A second way of bringing out the characteristically negative valence of "disruption" (as opposed to "change" *simpliciter*) is in terms of the action-theoretical character of our definition of social disruption, as impairments to individual and/or collective agency. To be disrupted is to be in a state of being unable to go on. And once disruption is understood in these agentic terms, then it turns out that – in many cases, at least – endorsing disruptions involves endorsing being put in a situation of not being able to go on with the only currently available way of meeting a key societal need. And that is *itself* rarely if ever something positive. Of course, a disruption to *others* might be something that one desires – that *they* are unable to go on. For example, we might be in favor of disruptions to the smooth operation of the fossil-fuel-based economy, but that is typically in the context of having alternatives available. And in that case, one is not desiring disruption *as such*. What one desires, quite rationally, is a shift from one way of coordinating the production of energy to alternative, preferable options. It is, of course, entirely rational to desire innovation and change, even when one foresees that the transformative process will have disruptive elements, as we have just discussed. Our point is only that a positive attitude towards a process that foreseeably involves disruptive elements is a matter of cost-benefit analysis or a determination of acceptable levels of *risk*, where the characterization as "costs" and "risks" highlights the fact that it would be preferable to bring about the desired transformation without the disruptive elements. In this sense, advocates for a "disruption of the fossil fuel industry" aren't typically advocating for a situation in which we would have no idea how to go on in providing basic needs for transport, heat, light, etc. would be met.

In short, the narrow construal foregrounds the element of disorientation in the process of disruption: not knowing how to continue, being unable to go on as before but also being at a loss as to how to move forward. This is a characteristic aspect of SDTs, but it does not deliver an all-encompassing evaluative verdict: the initial disorientation leaves open the possibility that, construed broadly and with a focus on their eventual outcomes, SDTs may also lead to moral growth and social progress.

3.4.3. Substantiality of a challenge is relative to availability of resources for responding

As we noted in Section 3.1, our definition aims to capture the *contextual* character of SDTs. In the present discussion, the key point is that whether the challenges generated by a technology are "substantial" in a given context also depends on the extent to which people are able to successfully meet those challenges. In general, while the challenges of SDTs can be substantial and pervasive, they are unlikely to persist indefinitely. The reason is simple: the narrowly negative valence of a disruption serves as a call to action. However, the success of strategies for coping with, adapting to, or overcoming challenges depends on the availability of resources for responding to SDTs, or what we call "coping resources". These include material, institutional, and epistemic resources to guide the implementation of a technology, to strengthen social and political resilience, or to counter powerful stakeholders who seek to appropriate the technology (see section 4). Such resources are unequally divided across communities and societies and amongst individuals, raising the prospect that the disruptive implications of technologies are experienced rather differently – and that they may raise substantial challenges for some, but not for others (we take these to be further aspects of 'differential disruption'). As an empirical matter, there may be some technologies for which no society is sufficiently well-prepared to respond quickly and effectively. But it remains important to avoid de-contextualized claims about the disruptiveness of specific technologies, and to emphasize that some individuals, communities, and societies are much more vulnerable to the challenges raised by SDTs, particularly in light of their respective resources for coping with, resisting, or adapting to disruptions. We will turn in the next section to the ethical concerns raised by this stratification of these resources.

4. Implications of the definition

We have argued that the defining properties of SDTs are their substantial challenges to normative orientation and joint action-coordination in a given context. In this section, we highlight the implications of this account for the ethics of socially disruptive technologies. As noted, our definition accommodates the differential ways in which SDTs can affect societies, depending on the resources available for normative orientation and action-coordination, and associated resources for resisting, adapting to, or coping with any disruption that does occur. These resources co-determine (1) whether the challenges generated by a technology are substantial enough for it to count as an SDT; (2) how problematic the *narrowly construed* disruptiveness is in the affected context (section 4.2); and (3) how beneficial or problematic the impact, *widely construed*, of the SDT ends up being. These are large, complex issues, requiring more analysis and discussion than we can undertake here. Our present aim is to highlight how our definition of SDTs helps to make ethical concerns with the potentially stratified availability of these resources more salient.

4.1. Stratification across societies: does a technology meet the substantially threshold?

We have defined SDTs as significantly meeting a threshold value of "substantiality" in challenging functionally key societal capacities for normative orientation or action-coordination, while emphasizing that this threshold value can vary from context to context. This has a relevant implication for ethical discourse on SDTs. Societies with particularly robust capacities for action-coordination and normative orientation may classify fewer potential technological changes as "socially disruptive." Their corresponding ethical discourse may be more open to the development and implementation of new technologies, and may have a narrower focus on a limited range of SDTs. Conversely, from the perspective of more vulnerable societies, more technologies may be experienced as disruptive and as cause for concern. This is important to keep in mind in the context of the complex issue of international regulatory governance.

In parallel with the label "high-risk" in the EU's AI Act, regulations regarding SDTs will have to find a way of accommodating differences in how substantially a technology threatens the societal capacities of different parties to any agreement. Our main suggestion here is to regard vulnerability – understood in terms of the availability of coping resources – as an integral contextual dimension in conceptualizing SDTs.

4.2. Three modes of response to disruptiveness: resistance, adaptation, and acquiescence

Our definition emphasizes that, even when their impact, *broadly construed*, is beneficial, the disruptive character of SDTs often lends them a negative dimension, in light of the burdensome and uncertain character intrinsic to the process of disruption (see section 3.4.2 above). Once again, the stratified availability of resources raises significant ethical concerns about conditions under which these narrowly negative implications might be aggravated. In what follows, we distinguish three characteristic ways in which agency can be exercised in responding to the (prospect of) narrowly negative implications of SDTs: resistance, adaptation, and acquiescing.

The first and most straightforward response is resistance, in which the prospect of disruption motivates efforts to halt or dramatically slow the implementation of an SDT. One source of the motivation to resist emerging technologies is the conviction that it would be *reckless* to implement a particular SDT in the absence of a clear, realistic plan for how the initial disruption of normative orientation and action-action coordination will be overcome and resolved. In a move familiar from one type of common appeal to the "precautionary principle", which foregrounds great and irreversible damage as a realistic possibility (e.g., Hopster, 2023), resistance is motivated by suspicions that the "move fast and break things" embrace of advocates of disruptive technologies betrays a dangerously naïve optimism about the ability of societies to restore vital societal functions. This can be seen in Hollywood actors striking against the introduction of GenAI in creative industries amidst a legal vacuum regarding the (re)use of, for example, actors' voices, or in the opposition to deploying blockchain technologies in "decentralized finance" without a clear understanding about how adequate new banking and currency arrangements would emerge (Jin, 2024).

The second, opposite, response focused on *adaptation* involves adopting a proactive stance to accommodate SDTs and address the uncertainties of their disruptive effects. This approach does not seek to counter the development and implementation of an SDT, but proposes ways in which new norms, procedures, regulatory safeguards, and conceptual frameworks can be feasibly and effectively implemented, in order to mitigate the SDT's narrowly negative implications. By transforming the conditions under which the technology is implemented – for example, by establishing legal frameworks or by rethinking central concepts that can form the basis of new normative orientations – proactive adaptations can mitigate the duration and severity of the disruptiveness. For instance, this mode of adaptation is the intention guiding the EU's AI Act, as a regulatory governance structure that can establish frameworks that will mitigate the disruptive effects of the implementation of numerous AI-related SDTs. However, it is not said that adaptation will also lead to overall improvement – to progress in the broad sense.

A third and more resigned approach is found in what we term "acquiescence," where the socially disruptive effects of the introduction of a technology are taken as given, perhaps even as inevitable, and a laissez-faire attitude is adopted to the ensuing social changes. From this perspective, the hope is often simply that, over time, preferences and norms may adjust to a new equilibrium, effectively normalizing the disruption without substantively addressing its challenges. An example of this can be found in responses to the challenges that social media pose to democratic institutions. Many societies' predominant response has been one of normalization: democratic norms and institutions have not substantially altered, while the use of extreme rhetoric and polarization

in the political arena has become increasingly accepted as the "new normal." While some may regard this as a failure to adequately respond to SDTs, others might maintain that a functional normative orientation had best develop organically, and the initial disorientation the disruption provokes will naturally recede.

In all these modes of addressing disruptiveness, even in the case of acquiescence, any success in coping in a way that actually mitigates the short-term consequences will depend on having societal resources and capabilities, including such factors as institutional support, levels of education, a strong civil society, communications infrastructure, level of security, *etc*. The inequalities in access to those resources can lead to significant differences in the degree to which societies are able to mitigate the (narrowly construed) disruptive impact of SDTs. Furthermore, the very nature of these resources may also be transformed by SDTs themselves, which calls for further scrutiny on the means for building a capacity for resilient response to SDTs.

4.3. Objection: Have we misplaced the locus of primary ethical concern?

It might be objected that our analysis fails to capture the most significant source of ethical concerns about SDTs: what is most troubling are not narrowly negative aspects intrinsic to the disruptive process, but the (broadly construed) negative impact and consequences SDTs may have. Instead of focusing on social and normative challenges, the objection runs, we should consider these eventual outcomes, and the harms and aggregate costs associated with them, as the defining ethical feature of social disruptions.

To illustrate the point, critics of our definition might argue that the reason nuclear weaponry counts as a "socially destructive technology" is its massively *destructive* potential, rather than the disorientation and joint-action orientation it provokes. The objection could be pressed further by referring to the fact that many SDTs – e.g. deep-sea mining, bioenergy with capture and storage, and even energy-intensive data centres – do not predominantly affect human beings, but the natural environment. It might seem that the disruptive potential of these technologies should not primarily be understood in terms of normative and social disorientation, but rather in terms of their environmental destruction. One might even claim that some of these technologies involve no real challenge about how to go on: from a clear-minded ethical point of view, one knows exactly what to do.

In response to this objection, there is one point we grant and one we resist. We grant that concrete and harmful impacts, including institutional failure and environmental destruction, are real and important aspects of what is colloquially referred to as "disruptions." There is no doubt that these aspects should take pride of place in an ethical assessment of SDTs. But we resist the suggestion that these destructive impacts are what *defines* SDTs – that is, that these are the primary features that make them *socially disruptive*. Nuclear technology can be highly *destructive* to human society, but its *disruptive* quality is not the destruction as such; rather, its disruptiveness arises from the normative and social challenges to which this destructive potential, or the destruction itself, gives rise. Indeed, not all destructive technologies are equally disruptive – a point that becomes clear when one considers that technology can sometimes be at its most destructive when it fits seamlessly into business as usual.

If the only ethical dimension of deep-sea mining were its clear-cut environmental costs, which could be uncontroversially quantified in a straightforward and widely endorsed cost-benefit analysis, then the technology would hardly constitute an SDT. But as a matter of fact, the technology is highly controversial in many respects: it raises vexed questions about the value of nature, international jurisdiction of the seabed, the power of corporations, and the nature of a just climate transition. It is in virtue of its accompanying normative and social challenges that the technology merits the label of an "SDT."

To make our response more precise, it is useful to distinguish between two theses about our definition:

Strong thesis: for SDTs, <u>all</u> of the urgency and gravity of the related ethical concerns can be fully explained in terms of their narrowly negative implications (i.e., as determined by challenges to normative orientation or action-coordination).

Weak thesis: for SDTs, <u>part</u> of the urgency and gravity of the related ethical concerns can be explained in terms of their narrowly negative implications (but potentially also in terms of ensuing harms, benefits, violations, etc.).

We favor the weak thesis. Accordingly, our definition identifies a set of criteria that all technologies deserving the label "SDT" should satisfy and that is jointly sufficient for calling something an SDT: they involve a substantial challenge to normative orientation or joint action-coordination. However, we grant that this need not be all there is to disruptiveness: a disruption may also involve other ethically important dynamics which should certainly not be overlooked, even though these do not *define* SDTs as a distinct class of technologies.

Having clarified our definitional framework and addressed potential objections, we now turn to the practical implications of understanding SDTs through the lens of functional disruption, implications that extend across multiple domains of technology governance and design.

5. Implications for a social ethics of disruptions

While our primary aim in this paper is conceptual, the account of SDTs developed in this paper – understanding disruption through the lens of challenges to normative orientation and joint action-coordination – carries significant implications for how we approach the governance of emerging technologies. By foregrounding the differential availability of coping resources across contexts, our framework reveals critical dimensions of technological justice that conventional approaches to technology assessment and regulation often obscure. In this penultimate section, we briefly trace these implications through several domains of practical application, providing an initial indication of how our theoretical framework reorients fundamental questions about risk assessment, regulatory design, and participatory governance.

5.1. Reconceptualizing risk assessment through a functionalist lens

Traditional risk assessment frameworks, as Aven and Hansson (Hansson & Aven, 2014) note, operate through a dual structure of factual classification and normative evaluation. Our definition of SDTs intervenes precisely at this juncture, proposing that the identification of a technology as "socially disruptive" constitutes not merely a descriptive categorization but a diagnosis of functional breakdown that demands distinctive forms of ethical scrutiny. Unlike conventional risk categories – where, for instance, the presence of chemical contaminants triggers engagement with toxicological expertise – the identification of an SDT signals a more complex evaluative challenge: the need to assess threats to fundamental social capacities whose restoration may require resources that are unequally distributed across global contexts.

This reconceptualization has significant implications for how we structure risk governance. When a technology threatens to undermine capacities for normative orientation – as deepfakes do for epistemic coordination, or as cultured meat does for cultural food practices – the relevant expertise cannot be confined to technical domains. Rather, assessment must engage with the socio-normative infrastructures that enable communities to navigate disruption. Crucially, our framework highlights that the same technology may pose radically different risk profiles across contexts, not merely because of varying exposure levels, but because communities do not always possess the same capacity to find new ways of going on in the aftermath of disruptions.

5.2. Regulatory frameworks and the challenge of contextual variation

The European Union's AI Act exemplifies both the promise and limitations of risk-based regulatory approaches when confronted with

SDTs. While the AI Act's tiered structure attempts to calibrate regulatory responses to risk levels, it presupposes a relatively stable understanding of what constitutes "high risk" – a presupposition that SDTs systematically undermine. Our analysis reveals why: SDTs are characterized precisely by their capacity to destabilize the normative frameworks within which risk assessments operate. The challenge is not simply that we lack mechanisms for responding to SDTs (though this is often true), but that the disruption extends to our very capacity to evaluate and coordinate responses (Marchiori et al., 2025).

This "second-order" disruptive character (Hopster & Maas, 2023) necessitates a fundamental rethinking of regulatory design. Rather than assuming universal risk thresholds, regulation of SDTs must incorporate mechanisms for recognizing and responding to contextual variation in disruptive impact and to be particularly sensitive to North-South differences. A technology that poses manageable challenges in contexts with robust institutional resources and high social capital may generate catastrophic disruption in contexts lacking such coping resources. Our framework thus points toward the need for what we might term "capacity-sensitive regulatory contextualism" – governance structures that can accommodate radical variation in how disruption manifests across different social contexts while maintaining coherent principles for evaluation and response.

In line with Lundgren's (2023b) model of the relationship between ethical guidelines, standards, and regulations, our theoretical work provides foundations for guidelines that must themselves be sensitive to the differential capacities of various contexts. This is not merely a matter of allowing for "local implementation" of universal principles, but of recognizing that the very meaning and impact of disruption varies with the availability of resources for maintaining or restoring social functionality.

5.3. From Value-Sensitive Design and to designing for disruption

The implications of our framework become particularly acute when we consider participatory and value-centered approaches to technology development, such as Value-Sensitive Design (VSD). These frameworks typically presuppose that relevant values can be identified through stakeholder engagement and then "designed into" technological systems. However, SDTs pose a fundamental challenge to this presupposition: they are defined precisely by their capacity to unsettle the normative orientations that would guide such value identification.

This highlights one of the challenges facing any attempt to "design for disruption": How can we design technologies to embody values when the technologies themselves threaten to transform or destabilize those very values? Our functionalist approach suggests a possible way forward by shifting focus from values to capacities. Rather than asking "what values should guide design?", we might ask "what resources and capabilities will communities need to navigate the disruptions this technology introduces?" (cf. Jacobs, 2020)

This reframing has concrete implications for design methodologies. In particular, it suggests the need for approaches that start out from an appreciation of differential access to coping resources across potential user communities and that recognize that "successful" design might mean enabling communities to develop their own responses, in light of iterative revisions of both proposed solutions and the shared understanding of what counts as success.

5.4. Social justice in the context of stratified of disruptive impacts

Perhaps nowhere are the ethical implications of our framework more significant than in questions of global technological justice. By highlighting how the substantiality of disruption depends on available resources for response, our analysis reveals a crucial dimension of technological inequality: communities with fewer resources for maintaining normative orientation and action-coordination face disproportionate disruption from the same technologies.

This stratification operates at multiple scales. Between nations, differences in institutional robustness, educational infrastructure, and social capital create vastly different capacities for navigating disruption. A facial recognition system that poses manageable challenges in contexts with strong data protection frameworks and active civil society organizations may fundamentally undermine social trust in contexts lacking such resources. Within nations, marginalized communities often lack access to the legal, social, and economic resources that enable effective responses to disruption (Lundgren et al., 2024).

Our framework thus points toward an expanded understanding of ethics of technology that goes beyond questions of access, privacy protections, or fair distribution to encompass, more broadly a "social ethics of technological disruption," with an emphasis on social justice and capacity building. The UN Declaration on the Right to Development (1986) provides an interesting resource for rethinking technology ethics along these lines, with an emphasis on institutional structures, material conditions, and social practices that support the development of adaptive capacity.

5.5. Toward anticipatory governance for functional resilience

The implications traced above converge on the need for new models of anticipatory governance that take seriously both the impact of SDTs on action-coordination and normative orientation and the stratified nature of adaptive capacity. Rather than focusing solely on preventing negative outcomes or maximizing benefits, governance frameworks for SDTs must attend to the preservation and enhancement of communities' capacities for normative orientation and joint action-coordination. This suggests, in turn, that the currently predominant focus on monitoring and compliance may need to be supplemented with a more proactive approach to investing in capacity-building and resilience planning. It also highlights the centrality of institutional precarity and geo-political vulnerability as key elements for adequately anticipating the effects of technology. Finally, it highlights the need for more attention to "transition ethics" and the development of normative frameworks that can guide action precisely during periods when established orientations have broken down and new equilibria have not yet emerged (cf. Hopster, 2025).

Our framework contributes to several ongoing debates about emerging SDTs. The rapid development of generative AI raises questions about epistemic coordination that our account of normative disorientation helps clarify (Al-kfairy et al., 2024; Öhman, 2024). The metaverse presents challenges to action-coordination across virtual and physical spaces that exemplify the functional breakdowns we identify (Brey, 2025; Dwivedi et al., 2022). Climate technologies like carbon removal systems disrupt existing frameworks for environmental responsibility in ways our contextual approach illuminates (McLaren & Corry, 2024; Taebi et al., 2023). This range of applications demonstrate how our definition provides analytical tools for assessing diverse technological disruptions.

6. Conclusion

Social disruptions challenge normative orientations and joint action-coordination, in ways that are disorienting, involve a loss of normative and social bearings, often create social friction, and may require risky, effortful resistance or costly material, technological, institutional or conceptual adaptations. We have argued that, narrowly construed, SDTs typically have a negative component: the disruptive process they provoke involves a loss of normative problem-solving capacity, as well as uncertainty and experiences of disorientation. Just how bad and lasting this disruption ends up being depends on the response and the available *coping resources*, which offers a key angle for the ethical evaluation of SDTs.

However, this narrowly negative valence of disruptiveness is compatible with SDTs bringing about significant improvements that

justify the risks, effort, and costs. SDTs may upend unjust relations of power and oppression, destabilize entrenched biases, dispel the grip of infertile concepts, and broaden horizons of normative and conceptual possibility. They may rupture hegemonic discourse and alter harmful patterns of human behavior. Social disruption may yield moral progress as well as moral regress, and oftentimes, there is substantial anticipatory uncertainty as to which of these outcomes is more plausible. This is the Janus-faced predicament that forms the starting point for ethical engagement with SDTs.

CRediT authorship contribution statement

Joel Anderson: Writing – review & editing, Writing – original draft, Conceptualization. **Jeroen Hopster:** Writing – review & editing, Writing – original draft, Conceptualization. **Björn Lundgren:** Writing – review & editing, Writing – original draft, Conceptualization.

Declaration of competing interest

Lundgren serves on the Ethics Advisory Board for the Estonian Centre of Excellence in Artificial Intelligence.

Data availability

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

References

Al-Kfairy, M., Mustafa, D., Kshetri, N., Insiew, M., & Alfandi, O. (2024). Ethical challenges and solutions of generative AI: An interdisciplinary perspective. *Informatics, 11*(3), 58. https://doi.org/10.3390/informatics11030058

Anderson, J. (2014). Regimes of autonomy. *Ethical Theory & Moral Practice, 17*(3), 355–368. https://doi.org/10.1007/s10677-013-9448-x

Baier, M. S., Berger, S., Kreuzer, T. M., Oberländer, A. M., & Röglinger, M. (2023). What makes digital technology? A categorization based on purpose. *Journal of the Association for Information Systems, 52*(4). https://doi.org/10.17705/1CAIS.05215

Baker, R. (2013). *Before bioethics: A history of American medical ethics from the colonial period to the bioethics revolution*. Oxford University Press.

Benagiano, G., Bastianelli, C., & Farris, M. (2007). Contraception: A social revolution. *The European Journal of Contraception and Reproductive Health Care, 12*(1), 3–12. https://doi.org/10.1080/13625180601012329

Biermann, F., Karcher, S., Möller, I., & Schwanholz, J. (2021). It is dangerous to normalize solar geoengineering research. *Nature, 595*, 30. https://doi.org/10.1038/d41586-021-01810-5

Boltanski, L., & Thévenot, L. (2006). *On justification: Economies of worth*. Princeton University Press.

Bovenkerk, B., & Boersma, K. (2023). Of mammoths and megalomaniacs. *Environmental Ethics, 45*(4), 381–402. https://doi.org/10.5840/enviroethics2023121582

Bower, J. L., & Christensen, C. M. (1995). Disruptive technologies: Catching the wave. *Harvard Business Review, 73*(1), 506–520.

Brey, P. (2017). Ethics of emerging technology. In S. O. Hansson (Ed.), *The ethics of technology: Methods and approaches* (pp. 175–191). Rowman & Littlefield.

Brey, P. (2025). *The metaverse: A critical assessment*. Switzerland: Springer Nature. https://doi.org/10.1007/978-3-031-93471-1

Brownsword, R. (2019). Law disrupted, law re-imagined, law re-invented. *Technology and Regulation, 1*, 10–30. https://doi.org/10.26116/techreg.2019.002

Carlsen, H., Dreborg, K. H., Godman, M., Hansson, S. O., Johansson, L., & Wikman-Svahn, P. (2010). Assessing socially disruptive technological change. *Technology in Society, 32*(3), 209–218. https://doi.org/10.1016/j.techsoc.2010.07.002

Choi, S., & Moon, M. J. (2023). Disruptive technologies and future societies: Perspectives and forecasts based on Q-methodology. *Futures, 145*, Article 103059. https://doi.org/10.1016/j.futures.2022.103059

Christensen, C. M. (1997). *The innovator's dilemma: The revolutionary book that will change the way you do business*. Harvard Business School Press.

Cocking, D., & van den Hoven, J. (2018). Evil online. *Blackwell Public Philosophy Series*. https://doi.org/10.1002/9781119471219

Congdon, M. (2024). *Moral articulation: On the development of new moral concepts*. Oxford University Press.

Danaher, J., & Sætra, H. S. (2023). Mechanisms of techno-moral change: A taxonomy and overview. *Ethical Theory & Moral Practice, 26*(1), 1–22. https://doi.org/10.1007/s10677-023-10397-x

Daub, A.. What tech calls thinking: An inquiry into the intellectual bedrock of silicon valley. *FSG Originals*. https://www.adriandaub.com/books/what-tech-calls.

De Proost, M., & Segers, S. (2023). We need to talk about disruption in bioethics: A commentary on Rueda, Pugh, and Savulescu. *Trends in Biotechnology*. https://doi.org/10.1016/j.tibtech.2022.12.00116

Dwivedi, Y. K., Hughes, L., Baabdullah, A. M., Ribeiro-Navarrete, S., Giannakis, M., Al-Debei, M. M., Dennehy, D., Metri, B., Buhalis, D., Cheung, C. M. K., Conboy, K., Doyle, R., Dubey, R., Dutot, V., Felix, R., Goyal, D. P., Gustafsson, A., Hinsch, C., Jebabli, I., … Wamba, S. F. (2022). Metaverse beyond the hype: Multidisciplinary perspectives on emerging challenges, opportunities, and agenda for research, practice and policy. *International Journal of Information Management, 66*, Article 102542. https://doi.org/10.1016/j.ijinfomgt.2022.102542

Fallis, D. (2021). The epistemic threat of deepfakes. *Philosophy & Technology, 34*(4), 623–643. https://doi.org/10.1007/s13347-020-00419-2

Feenberg, A. (2017). *Technosystem: The social life of reason*. Harvard University Press.

Frank, L. E., Hermann, J., Kavege, L., & Puzio, A. (2023). Ectogestative technology and the beginning of life. In *Ethics of socially disruptive technologies* (pp. 113–140). Open Book Publishers. https://doi.org/10.11647/obp.0366.05.

Fritz, L., Baum, C. M., Low, S., & Sovacool, B. K. (2024). Public engagement for inclusive and sustainable governance of climate interventions. *Nature Communications, 15*(1), 4168: 1–17. https://doi.org/10.1038/s41467-024-48510-y

Fukuyama, F. (2017). *The great disruption*. Profile Books.

Gilbert, M. (2014). *Joint commitment: How we make the social world*. Oxford University Press.

Giovanola, B. (2023). Justice, emotions, socially disruptive technologies. *Critical Review of International Social and Political Philosophy, 26*(1), 104–119. https://doi.org/10.1080/13698230.2021.2001311

Groff-Vindman, C. S., Trump, B. D., Cummings, C. L., Klasa, K., & Linkov, I. (2025). The convergence of AI and synthetic biology: The looming deluge. *npj Biomedical Innovations, 2*, 20. https://doi.org/10.1038/s44385-025-00021-1

Habgood-Coote, J. (2023). Deepfakes and the epistemic apocalypse. *Synthese, 201*, 103. https://doi.org/10.1007/s11229-023-04097-3

Hansson, S. O., & Aven, T. (2014). Is Risk Analysis Scientific? *Risk Analysis, 34*(7), 1173–1183. https://doi.org/10.1111/risa.12230.

Harbin, A. (2016). *Disorientation and moral life*. Oxford University Press. https://doi.org/10.1093/acprof:oso/9780190277390.001.0001

Haslanger, S. (2020). Going on, not in the same way. In A. Burgess, H. Cappelen, & D. Plunkett (Eds.), *Conceptual engineering and conceptual ethics* (pp. 230–260). Oxford University Press.

Hayward, C. R. (2020). Disruption: What is it good for? *The Journal of Politics, 82*(2), 448–459. https://doi.org/10.1086/706766

Heath, J. (2006). The benefits of cooperation. *Philosophy & Public Affairs, 34*(4), 313–351. https://doi.org/10.1111/j.1088-4963.2006.00073.x

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

Hermann, J. (2025). Moral certainty, deep disagreement, and disruption. *Synthese, 205* (3), 1–18. https://doi.org/10.1007/s11229-024-04843-1

Hopster, J. K. G. (2021a). What are socially disruptive technologies? *Technology in Society, 67*, Article 101750. https://doi.org/10.1016/j.techsoc.2021.101750

Hopster, J. K. G. (2021b). Mutual affordances: The dynamics between social media and populism. *Media, Culture & Society, 43*(3), 551–560. https://doi.org/10.1177/0163443720957889

Hopster, J. K. G. (2023). Climate uncertainty, real possibilities and the precautionary principle. *Erkenntnis, 88*(6), 2431–2447. https://doi.org/10.1007/s10670-021-00475-w

Hopster, J. K. G. (2025). The weight of legitimate expectations in a just climate transition. *Journal of Applied Philosophy*. https://doi.org/10.1111/japp.12775

Hopster, J. K. G., Arora, C., Blunden, C., Eriksen, C., Frank, L. E., Hermann, J. S., Klenk, M. B. O. T., O'Neill, E., & Steinert, S. (2022). Pistols, pills, pork and ploughs: The structure of technomoral revolutions. *Inquiry*, 1–33. https://doi.org/10.1080/0020174X.2022.2090434

Hopster, J., Brey, P., Klenk, M., Löhr, G., Marchiori, S., Lundgren, B., & Scharp, K. (2023). Conceptual disruption and the ethics of technology. In *Ethics of socially disruptive technologies* (pp. 141–162). Open Book Publishers. https://doi.org/10.11647/obp.0366.06.

Hopster, J. K. G., & Löhr, G. (2024). *Conceptual engineering and philosophy of technology: Amelioration or adaptation? Philosophy and Technology, 37* p. 134). https://doi.org/10.1007/s13347-024-00810-3

Hopster, J. K. G., & Maas, M. (2023). The technology triad: Disruptive AI, regulatory gaps, and value change. *AI and Ethics*. https://doi.org/10.1007/s43681-023-00327-z

Jacobs, N. (2020). Capability sensitive design for health and wellbeing technologies. *Science and Engineering Ethics, 26*(6), 3363–3391. https://doi.org/10.1007/s11948-020-00275-5

Jaeggi, R. (2018). Resistance to the perpetual danger of relapse: Moral progress and social change. In A. Allen, & E. Mendieta (Eds.), *From alienation to forms of life: The critical theory of Rahel Jaeggi* (pp. 15–40). Pennsylvania State University Press.

Jin, S. (2024). The paradigm logic of blockchain governance. *Technology in Society, 78*, Article 102681. https://doi.org/10.1016/j.techsoc.2024.102681

Lane, B. (2025). Socially disruptive technologies, moral progress, and rule following. *Philosophy & Technology, 38*(2), 1–30. https://doi.org/10.1007/s13347-025-00797-x

Liu, H. Y., Maas, M., Danaher, J., Scarcella, L., Lexer, M., & Van Rompaey, L. (2020). Artificial intelligence and legal disruption: A new model for analysis. *Law, Innovation and Technology, 12*(2), 205–258. https://doi.org/10.1080/17579961.2020.1815402

Löhr, G. (2023). Conceptual disruption and 21st century technologies: A framework. *Technology in Society, 74*, Article 102327. https://doi.org/10.1016/j.techsoc.2023.102327

Lundgren, B. (2021). Improving on and assessing ethical guidelines for digital tracking and tracing systems for pandemics. *Ethics and Information Technology, 23* (Supplement 1), 139–142. https://doi.org/10.1007/s10676-020-09561-z

Lundgren, B. (2023a). Ethical requirements for digital systems for contact tracing in pandemics: A solution to the contextual limits of ethical guidelines. In K. Macnish, & A. Henschke (Eds.), *The ethics of surveillance in times of emergency* (pp. 169–185). Oxford University Press. https://doi.org/10.1093/oso/9780192864918.003.0011.

Lundgren, B. (2023b). Defense of ethical guidelines. *AI Ethics, 3*, 1013–1020. https://doi.org/10.1007/s43681-022-00244-7

Lundgren, B. (2025). Can deepfakes violate an individual's moral right to privacy? *Ethical Theory & Moral Practice*. https://doi.org/10.1007/s10677-025-10514-y

Lundgren, B., Catena, E., Robertson, I., Hellrigel-Holderbaum, M., Jaja, I. R., & Dung, L. (2024). On the need for a global AI ethics. *Journal of Global Ethics, 20*(3), 330–342. https://doi.org/10.1080/17449626.2024.2425366

Maas, M. M. (2022). Aligning AI regulation to sociotechnical change. In J. Bullock, B. Zhang, Y. Chen, J. Himmelreich, M. Young, A. Korinek, & V. Hudson (Eds.), *Oxford handbook on AI governance*. Oxford University Press. https://doi.org/10.2139/ssrn.3871635.

Marchant, G. E. (2011). The growing gap between emerging technologies and the law. In *The growing gap between emerging technologies and legal-ethical oversight: The pacing problem* (pp. 19–33). Springer.

Marchiori, S., Hopster, J. K. G., Puzio, A., Riemsdijk, M. B. V., Kraaijeveld, S. R., Lundgren, B., Viehoff, J., & Frank, L. E. (2025). A social disruptiveness-based approach to AI governance: Complementing the risk-based approach of the AI act. *Science and Engineering Ethics, 31*(5), 25. https://doi.org/10.1007/s11948-025-00545-0

Marchiori, S., & Scharp, K. (2024). What is conceptual disruption? *Ethics and Information Technology, 26*(1). https://doi.org/10.1007/s10676-024-09748-8

Mason, R. (2016). The metaphysics of social kinds. *Philosophy Compass, 11*(12), 841–850. https://doi.org/10.1111/phc3.12381

McLaren, D., & Corry, O. (2024). The politics of net zero: Techno-optimism and the limits of carbon management. *Environmental Politics, 33*(2), 189–208. https://doi.org/10.1080/09644016.2023.2169800

Mendly-Zambo, Z., Powell, L. J., & Newman, L. L. (2021). Dairy 3.0: Cellular agriculture and the future of milk. *Food, Culture and Society, 24*(5), 675–693. https://doi.org/10.1080/15528014.2021.1888411

Nickel, P. J. (2020). Disruptive innovation and moral uncertainty. *NanoEthics, 14*(3), 259–269. https://doi.org/10.1007/s11569-020-00375-3

Nickel, P. J., Kudina, O., & Van de Poel, I. (2022). Moral uncertainty in technomoral change: Bridging the explanatory gap. *Perspectives in Science*. https://doi.org/10.1162/posc_a_00414

O'Neill, E. (2022). Contextual integrity as a general tool for evaluating technological change. *Philosophy and Technology, 35*, 79. https://doi.org/10.1007/s13347-022-00574-8

Öhman, C. (2024). The epistemic coup: How generative AI is transforming knowledge production. *Philosophy & Technology, 37*(2). https://doi.org/10.1007/s13347-024-00722-2

Ostrom, E. (2000). Collective action and the evolution of social norms. *The Journal of Economic Perspectives, 14*(3), 137–158. https://doi.org/10.1257/jep.14.3.137

Ostrom, E. (2010). Beyond markets and states: Polycentric governance of complex economic systems. *The American Economic Review, 100*(3), 641–672. https://doi.org/10.1257/aer.100.3.641

Otto, K., & Köhler, A. (Eds.). (2018). *Trust in media and journalism: Empirical perspectives on ethics, norms, impacts and populism in Europe*. Springer.

Possati, L. M. (2023). Ethics of quantum computing: An outline. *Philosophy & Technology, 36*, 48. https://doi.org/10.1007/s13347-023-00651-6

Rini, R. (2020). Deepfakes and the Epistemic. *Backstop, 20*(24), 16.

Rotolo, D., Hicks, D., & Martin, B. (2015). What is an emerging technology? *Research Policy, 44*(10), 1827–1843. https://doi.org/10.1016/j.respol.2015.06.006

Rueda, J., Pugh, J., & Savulescu, J. (2022). The morally disruptive future of reprogenetic enhancement technologies. *Trends in Biotechnology*. https://doi.org/10.1016/j.tibtech.2022.10.007

Sauer, H. (2024). *The invention of good and evil: A world history of morality*. Oxford University Press.

Schatzberg, E. (2018). *Technology: Critical history of a concept*. University of Chicago Press.

Schuelke-Leech, B. (2018). A model for understanding the orders of magnitude of disruptive technologies. *Technological Forecasting and Social Change, 129*, 261–274. https://doi.org/10.1016/j.techfore.2017.09.033

Scott, W. R. (2014). *Institutions and organizations: Ideas, interests, and identities*. SAGE.

Scully, J. L. (2019). Deaf identities in disability studies. In *Routledge handbook of disability studies* (pp. 145–157). Routledge.

Siegrist, M., Michel, F., & Hartmann, C. (2024). The shift from meat to plant-based proteins: Consumers and public policy. *Current Opinion in Food Science, 58*, Article 101182. https://doi.org/10.1016/j.cofs.2024.101182

Skyrms, B. (2010). *Signals: Evolution, learning, and information*. Oxford University Press.

Swierstra, T. (2013). Nanotechnology and technomoral change. *Etica e Politica/Ethics and Politics, XV*(1), 200–219.

Taebi, B., Lenzi, D., Buhr, L., Claassen, K., Gerola, A., Hofbauer, B., Paiusco, E., & Rijssenbeek, J. (2023). Climate engineering and the future of justice. In *Ethics of socially disruptive technologies* (pp. 83–112). Open Book Publishers. https://doi.org/10.11647/obp.0366.04.

Van de Poel, I., Hermann, J., Hopster, J., Lenzi, D., Nyholm, S., Taebi, B., & Ziliotti, E. (2023). *Ethics of socially disruptive technologies: An introduction*. Open Book Publishers. https://doi.org/10.11647/obp.0366

Van Grunsven, J., & Marin, L. (2024). Technosocial disruption, enactivism, & social media: On the overlooked risks of teenage cancel culture. *Technology in Society, 78*, Article 102602. https://doi.org/10.1016/j.techsoc.2024.102602

Whicker, M. L., & Kronenfeld, J. J. (1986). Men and women together: The impact of birth control technology on male-female relationships. *International Journal of Sociology of the Family, 16*(1), 61–81.

Zakaria, F. (2024). *Age of revolutions: Progress and backlash from 1600 to the present*. W. W. Norton.