Towards a research ethics of real-world experimentation with emerging technology

Mollen, Joost

DOI
10.1016/j.jrt.2024.100098

Publication date
2024

Document Version
Final published version

Published in
Journal of Responsible Technology

Journal of Responsible Technology 20 (2024) 100098

Contents lists available at ScienceDirect

Journal of Responsible Technology

journal homepage: www.sciencedirect.com/journal/journal-of-responsible-technology

Research Article

Towards a research ethics of real-world experimentation with emerging technology

Joost Mollen

Faculty of Technology, Policy and Management, Delft University of Technology (NL), Jaffalaan 5, 2628BX, Delft, The Netherlands

ARTICLE INFO

Keywords:
Real-world research
Field experimentation
Research ethics
Unequal treatment
Regulatory arbitrage
Emerging technology

ABSTRACT

Testing emerging technologies, such as autonomous vehicles, predictive crime analytics, and smart city interventions under real-world conditions is an important strategy for robust and responsible technology development. However, the moral responsibilities of researchers towards the public when conducting such real-world experiments are often left unaddressed and unregulated. This article argues that there are problematic inconsistencies in research ethics demands and protections across different categories of research and development with emerging digital technologies. This differential treatment is problematic since there are no meaningful differences to justify it, and it creates the possibility of regulatory evasion at the cost of populations' due protection. Hence, I argue that this differential treatment should be amended by harmonizing research ethics demands. In doing so, this paper contributes to several ongoing scholarly debates on the limits of current research ethics guidelines and protocols in the face of novel technologies and research formats.

1. Introduction

Testing emerging technologies under real-world conditions has emerged as a distinctive data-driven strategy to address various societal and innovation challenges in recent decades. This 'real-world' research and development is understood as an important strategy to bridge performance in controlled laboratory environments to eventual successful real-world deployment and to study to what degree a particular emerging technology can be leveraged to solve a particular social issue (Ansell & Bartenberger, 2016; Pfotenhauer et al., 2022). Examples range from online experiments such as algorithmic A/B tests with emotions, romantic relations, careers, and wages on social media or online worker platforms (Grimmelman, 2015; Polonioli et al., 2023; Rahman et al., 2023; Wood, 2020) to experimentation in (urban) physical environments, with self-driving cars on public roads (DeArman, 2019; Stilgoe, 2020), predictive policing and crowd control technologies in nightlife streets and neighborhoods (Amnesty, 2020; Galič, 2019; Susser, 2021), smart homes (Taylor, 2021; Mollen, 2023) and smart city interventions (Zimmermann, 2023).

While real-world research might benefit the development of responsible emerging technologies (Colonna, 2023) or help solve social challenges (Ansell & Bartenberger, 2016), attention should also be paid to conducting these experiments *responsibly*. Since real-world experiments operate closely to people's daily lives or environment and actively intervene within them, they potentially cause undue influence, impact, or harm to persons who become (un)knowingly or (un)desirably involved. For example, Colonna notes that artificial intelligence "that is tested in real-world conditions" .. can present "risks to individual's health, safety and fundamental rights, as well as broader societal concerns" (Colonna, 2023, p.28). For example, testing experimental AI facial recognition systems that turn out to be biased can harm individuals or groups of people through discrimination and non-equal treatment, but also, as Smuha points out, can cause broader societal harms such as "a higher interest to live in a society that does not discriminate against people based on their skin color and that treats its citizens equally" (Smuha, 2021, p.6).

However, while testing emerging technologies under real-world conditions raises various ethical issues, they are often not bound by research ethics regulations like other forms of research. Significant discrepancies exist between the regulatory demands placed upon different research and development categories. Often, these discrepancies follow the scope of institutional borders and funding-related obligations. However, they result in unequal access and exposure to protections and standards for different research participants or bystanders.

In this paper, I argue that this differential treatment is morally problematic because no meaningful conceptual differences between currently ethically regulated and unregulated research justify this

E-mail address: j.k.mollen@tudelft.nl.

https://doi.org/10.1016/j.jrt.2024.100098

Available online 21 October 2024

treatment. Additionally, due to the collaborative nature of modern research practices, this differential treatment creates the possibility of research ethics regulatory evasion, which comes at the expense of those whom these demands aim to help protect. Consequently, I argue that these discrepancies should be amended by harmonizing research ethics demands across research practices.

In doing so, this paper contributes to (1) the increased awareness about the research ethical dimensions of real-world research with emerging technologies (Taylor, 2021; Polonioli et al., 2023; Rahman et al., 2023; Zimmerman, 2023), (2) to various ongoing debates on the necessity of research ethics reforms regarding real-world research within AI and data science and social and political science, and (3) to ethics-by-design approaches that have called for the inclusion of ethical considerations in the design and development process of new technological systems and devices by drawing attention to a category of research and development that need 'research' ethics-by-design (Dainow & Brey, 2021).

This paper proceeds as follows. In the first section, I draw from two ongoing debates on research ethics reforms regarding real-world research within AI and data science and social and political science to ground my call for research ethics harmonization within a more extensive debate on the current shortcomings of research ethics guidelines. In the second section, I extend the focus to discuss the increasing attention to the lack of ethical standards and regulations in various forms of ethically unregulated real-world research. In the third section, I argue that the current differential treatment is problematic since no meaningful differences justify this treatment. In the fourth section, I argue that this differential treatment creates the possibility of 'research ethics' regulatory evasion due to the collaborative nature of modern research practices. Finally, I briefly discuss several avenues and challenges to resolving this issue.

2. The growing needs and demands for real-world research ethics

In this section, I describe the moral salience of real-world research and discuss two ongoing debates on the need for research ethics reforms in field experimentation, specifically from the perspective of AI and data science and social and political experimentation. By connecting these two debates, I ground my later argument about harmonizing research ethics principles and practices with currently unregulated research in a broader ongoing academic debate about research ethics reforms.

Real-world or field research, as the name implies, involves conducting research and testing under 'real-world conditions.' It differs from laboratory or controlled research since no experimental controls are placed on the research environment. We can distinguish between observational and experimental real-world research. The first studies phenomena that arise naturally; the latter actively brings about the phenomena studied through active intervention[1]. A theoretical example of the former would be using digital technology to capture location data to measure crowd density; an example of the latter would be studying how the use of various phrases on public digital billboards could influence crowd density. Both observational and experimental real-world research raises ethical concerns. In both cases, researchers place themselves or whatever they study in a participant's or community's daily lives or environment (Teele, 2014). Real-world experimentation adds extra moral weight due to being responsible for creating the data they observe (Grimmelman, 2015). By intentionally altering the environment, they can bring about undesirable, unintended, and unforeseen consequences caused by the research intervention (Teele, 2014, p. 119).

Scholars increasingly call for research ethics reforms regarding the design or conduct of real-world research. One domain in which this debate is prominent is within AI and data science. Scholars have discussed the ethical and regulatory challenges for researchers and institutional review boards regarding social media and online data research (Moreno et al., 2013; Raymond, 2019; Vitak et al., 2016, p. 277). Metcalf and Crawford have pointed out the missing focus on human subjects in big data science (2016). While databases with subject data can be infinitely re-combined to formulate pictures that are much more invasive than the initial study might be, issues of human subject research ethics often stay out of focus due to current formulations of what constitutes human subject research (Metcalf & Crawford, 2016). This point has also been made in a recent report by the Ada Lovelace Institute, in which the authors write that "the current role, scope, and function of most academic and corporate RECs [*red.* research ethics committees] are insufficient for the myriad of ethical challenges that AI and data science research can pose. For example, the scope of REC reviews is traditionally only on research involving human subjects. This means that the many AI and data science projects that are not considered a form of direct intervention in the body or life of an individual human subject are exempt from many research ethics review processes" (2022, p. 8). In light of these limitations, Resseguier and Ufert have argued in favor of adapting current research ethics standards and mechanisms to better asses scientific AI field research (Resseguier & Ufert, 2023). These discussions point to AI field research challenging existing research ethics and the need for research ethics redesign.

The ethics of field experimentation has also been the topic of a recent political and social science debate. In recent decades, these disciplines have increasingly moved toward field experimentation (experimental interventions outside controlled laboratory environments) as a prominent research methodology. However, this move beyond the lab has prompted questions of ethics and growing calls by scholars that these social and political field experiments are insufficiently ethically regulated (Humphreys, 2015; Teele, 2014; Desposato, 2018; Beerbohm et al., 2020; MacKay, 2018; McDermott & Hatemi, 2020; Phillips, 2021; Whitfield, 2019). For example, McDermott and Hatemi note about political field experiments that "we have somehow entered into a Wild West where anything goes when it takes place in the public sphere in large populations, while small controlled laboratory experiments must follow established guidelines" (McDermott & Hatemi, 2020, p. 30019).

Researchers have also pointed out that these political field experiments might bring about unique or different ethical concerns vis-à-vis other types of research. For example, Beerbohm and colleagues have argued that political field experiments may undermine political equality between citizens (Beerbohm et al., 2020). McDermott and Hatemi have argued that social and political field experiments may harm entire societies, which current research ethics frameworks cannot account for since they focus on transgressions against the individual (McDermott & Hatemi, 2020). Instead, they have suggested 'respect for societies' as an action-guiding and protecting principle in the design and execution of

[1] This distinction between observational and interventional research is not necessarily sharp. One reason is conceptual unclarity about what qualifies as a manipulation or intervention in the research context. For example, while ethnographic research is considered observational research, the presence of researchers within these communities can influence the observed behavior. One way to resolve this, is to fall back on the researcher's intentions. As Teele puts it, this difference concerns the degree to which "the researcher *purposively* manipulates the research context in some way" (Teele, 2014, p. 118; italics my own).

social and political experiments within society.[2] Additionally, Whitfield has argued for developing a separate research ethics for political field experiments (Whitfield, 2019). He argues that political field experimentation involving human subjects shares the capacity to commit interpersonal and diffuse wrongs with biomedical research but that only political field experiments may bring about wrongs against 'collectives'. Such collective wrongs undermine the decision-making capacity and rights of *groups* rather than individuals, particularly by undermining underlying values such as sovereignty ("the right of states to noninterference in their internal affairs"), subsidiarity ("the rights and authorities of states devolve to substate units and organizations") and association ("the rights of individuals to form associations") (Whitfield, 2019, p.533). Specific examples of political field experiments that might undermine, for example, sovereignty are those that "involve intentionally violating foreign laws" (Whitfield, 2019, p.533), as he claims was the case in (Fried et al., 2010).[3][4]

While the current debates within scientific research have rightfully drawn attention to both the ethics of field research and that current research ethics guidelines do not capture the full range of moral issues various research fields encounter, they leave a significant gap: they have been primarily focused on the expansion or adaption of research ethics to *scientific* research outside the laboratory. However, much 'non-scientific' research currently lacks regulatory ethical demands. Calls to expand and reform research ethics practices and protocols have only limitedly been extended to this research category. I will expand on this in the next section.

3. Scientific research ethics exceptionalism

Scholars have increasingly pointed out that research as a domain of human activity seems much more stringently regulated than many seemingly identical activities. The fact that scientists who want to conduct interviews for research purposes need to get approval from an ethical review board while journalists do not is one example of this. However, we find such examples anywhere where research and non-research activities overlap, whether fishing, urban planning, traffic, policy-making, sports, or business (Hansson, 2011). It seems that a vast range of ethically salient activities can be conducted as both (part of) research or outside a research setting. In turn, whether or not something is understood to be research determines the ethical demands placed on this activity. Research activities are thus seemingly treated as 'exceptional' (Wilson & Hunter, 2010). They are subject to higher ethical demands than similar human activities that are not labeled as research (Wilson & Hunter, 2010). This difference is particularly noticeable when it comes to human subject research. Scholars have disagreed over whether this difference is justified (Hansson, 2011; Wilson & Hunter, 2010). This problem is what Hansson has called the *boundary problem of research ethics*: what exactly – if anything – justifies this differential treatment of ethical demands (Hansson, 2011)?

However, one shortcoming of these accounts is that not *all* research is treated exceptionally. Different kinds of research are not held to the same moral standards (Miller, 2010; Moffat, 2010). This boundary of ethical demands that Hansson mentions does not run 'around' research; it cuts right through it. While *scientific* or *governmental-funded* research is often held to clear (but perhaps imperfect) research ethical standards and protocols, the same cannot be said for research conducted by many public and private parties. For example, Poloni and colleagues observe that while "protection protocols have become the norm in medical research and the social and behavioral sciences," … "the use of human subjects in research that is not federally or publicly funded—such as in the case of privately funded A/B testing, often affecting millions of potentially unaware people—has remained unregulated" (Polonioli et al., 2023, p. 669). Similarly, in the context of corporate experimentation on online worker platforms, Rahman writes: "the problem is not experimentation in itself, which can be useful to help companies make data-driven decisions. It is that most do not have any internal or external mechanisms to ensure that experiments are clearly beneficial to their users, as well as themselves. Countries also lack strong regulatory frameworks to govern how organizations use online experiments and the spillover effects they can have. Without guardrails, the consequences of unregulated experimentation can be disastrous for everyone" (Rahman, 2024). This has been an apparent issue for a significant period already. Already in 2002, Calo wrote that: "today, any academic researcher who would conduct experiments involving people is obligated to comply with robust ethical principles and guidelines for the protection of human subjects, even if the purpose of the experiment is to benefit those people or society… But a private company that would conduct experiments involving thousands of consumers using the same basic techniques, facilities, and personnel faces no such obligations, even where the purpose is to profit at the expense of the research subject" (Calo, 2013, p.101). Privately conducted research, thus, is not subject to the same research ethics regulation as scientific or publicly-funded research (Moffat, 2010).

To clarify, this is not merely a phenomenon of *corporate* research but also includes research conducted by governmental or public parties such as law enforcement agencies that might lack institutional obligations regarding research ethics standards. In the context of testing new technologies for migration management in border zones, Molner writes that "all this experimentation occurs in a space that is largely unregulated, with weak oversight and governance mechanisms, driven by the private sector innovation" (Molnar, 2020, p. 34). Alternatively, in 2020, Amnesty International called on the Dutch government to end dangerous police experiments with mass surveillance, citing human rights abuses (Amnesty International, 2020). In their report, Amnesty outlines the dangers of 'predictive policing,' a method that uses mathematical models to assess the likelihood of a criminal offense being committed by a specific individual or at a particular location (Amnesty, 2020; Susser, 2021). Amnesty writes: "a comprehensive policy and legal framework for regulation and oversight is yet to be introduced. Meanwhile, the police are running several experimental predictive policing projects under the premise that the existing legal framework sufficiently regulates their use of algorithmic models and big data predictions" (2020, p.14). These accounts indicate that various corporate and public research forms are often unregulated or comparatively less regulated than their scientific or publicly funded counterparts.

This is not to say these practices are necessarily bereft of ethical

[2] It remains unclear, however, in McDermott and Hatemi's argument what this principle entails, what would satisfy it, and to what extent it would cover a theoretical deficiency, for example, why 'respect for persons' as a principle is distinct from the societal cumulation of 'respect of persons'. However, McDermott and Hatemi invoke examples that exemplify individual moral concerns, such as not consenting to research participation, or due to the experiment changing features of the world that individuals generally have rights against, such as the non-target population being exposed to increased risks to their welfare. Additionally, it remains unclear how their principle of respect for society relates to those conditions where an overall society may be considered unjust, and the purpose of the research intervention is to (help) solve said injustices.

[3] Whitfield mentions he does not wish to suggest that "lab experiments, semi-, or nonexperimental methods cannot in principle risk similar impacts" as political field experiments (2019, p.536). This somewhat undermines his appeal for a separate research ethics of political field experiments. Instead, it seems rather to suggest that research ethics, in general, should be more cognizant of the specific impacts Withfield draws attention to, which may be caused by political field experiments but also by other kinds of research.

[4] Mackay (2023) has criticized these accounts, noting that, while valuable, they are building on a limited and outdated image of research ethics, for example, by only mentioning the Belmont Report (1978). He notes that: "discussions of the ethics of clinical research have moved beyond this, refining interpretations of the principles and applications found in Belmont, contesting these interpretations, and developing new concepts and principles to evaluate clinical research protocols" (Mackay, 2023, p.3).

regulations altogether. Many companies or public institutions have integrated ethics into their operations through ethics committees, codes of ethics, or internal guidelines. However, at the core, these practices often amount to self-regulation. Its value notwithstanding, what separates these practices from ethics regulation in scientific research on at least three counts is that (1) it is externally imposed, (2) it is mandatory, and (3) it holds actual sway over whether the research is conducted or can continue. For the sake of brevity going forward, I will refer to these two categories as ethically regulated research and ethically unregulated research when describing those kinds of research bound in some sense by externally imposed and mandatory ethical regulations and those that are not.

There have been some calls in the literature to extend ethical guidelines to various areas of ethically unregulated research. Recently, various scholars have discussed the ethics of corporate social media A/B testing and have called to extend current ethical regulations – such as institutional review boards – to these practices (Benbunan-Fich, 2017; Grimmelman, 2015; Jouhki et al., 2016; Kramer et al., 2014; Polonioli et al., 2023; Wood, 2020). Svensson and Hansson have discussed extending research ethics guidelines to traffic experiments (Svensson & Hansson, 2007). Zimmerman has used research ethics guidelines from psychology to analyze experimental smart city interventions (Zimmermann, 2023). Taylor has argued for research ethics in urban experimentation, arguing that under corporate technological field experimentation, research subject populations suffer a problematic power asymmetry and are "disempowered with respect to knowledge, understanding and agency" (Taylor, 2021, p. 1909).

The point of this paper differs from those of these scholars in at least two meaningful ways. First, the focus of this paper does not focus on a specific form of testing under real-world conditions, such as A/B tests. Instead, it takes a broader perspective to include various forms of research and development with technology under real-world conditions to make a broader case for harmonization. Second, this body of work has not significantly engaged with the idea that this differential treatment can be justified if meaningful differences exist. While authors such as Hansson (2011) and Wilson and Hunter (2010) have questioned whether scientific research needs more stringent ethical regulation compared to non-research activities of similar risk, there has been limited scholarly attention to examining the justification of different ethical demands between scientific and non-scientific research practices. Hence, in the next section, I examine this idea. I will argue that there are no meaningful conceptual differences between these two categories of research that justify the current different regulatory demands.

4. Arguments against meaningful conceptual difference

In this section, I will argue that this treatment of ethical regulation for different forms of research is problematic since no meaningful conceptual differences exist between currently ethically unregulated and ethically regulated research. I will examine four possible reasons that might justify this differential treatment. These are (1) the ability or potential to be harmful, (2) their environmental research conditions, (3) the goal of the research, and (4) the role of the researcher. I will subsequently reject these as sound justificatory reasons.

First, an argument to justify the current differential treatment could be to claim that all those forms of currently unregulated research are less or not risky than those currently regulated research (Wilson & Hunter, 2010).[5] Introducing ethical regulation would, in that case, amount to disproportionate over-regulation of research that is not (likely to) cause any harm. However, this position is unconvincing. While it may be true that *some* unregulated research with human subjects is not risky, there are plenty of examples of research that are risky or have caused harm, such as an experimental self-driving vehicle operated by Uber killing a pedestrian (Stilgoe, 2020), the Dutch police violated human rights with their predictive policing experiments (Amnesty, 2020), and Uber reshaping the sense of autonomy of platform workers (Rahman et al., 2023). In other words, independent of their current ethical regulatory status, ethically unregulated real-world experiments with digital technology can carry some risk. This is the same for currently ethically regulated research; some are more risky than others. We find examples of risky and non-risky research in ethically regulated and unregulated research.[6] Thus, there should be no difference between the regulatory status of these two categories of research.

To clarify, I do not claim that ethically unregulated research is *more* risky than ethically regulated research based on its regulatory status or that regulation will necessarily prevent these abuses. As Wilson and Hunter point out, while cases of research risk, abuse, or harm "do provide prima facie evidence that unregulated research *can* be abused […] they do not show that regulation *will* prevent these abuses" (Wilson & Hunter, 2010, p.49; italics my own). Merely having ethical regulation does not mean that researchers will keep to it. In order to make this strong argument for ethical regulation of currently unregulated research, I would either have to demonstrate that (1) there is a direct relation between an absence of research ethics regulation and concrete research abuses or that (2) research ethics guidelines do minimize risks and realize participant safety. However, these are empirical claims, and not only are they outside the scope of this paper, but little empirical evidence exists to support them (Bean, 2010). Instead, the argument above has been limited to claiming that no apparent difference exists in the potential for harm between ethically unregulated and regulated forms of research.

Second, there are no moral differences between the environmental conditions under which this research is conducted since both regulated and unregulated research can be conducted under similar conditions, namely real-world conditions. There might be strong arguments that moral obligations can be waived in given research contexts, such as that studying behavior in the public sphere does not require consent (Spicker, 2011). However, this is an argument about prioritizing or waiving two competing rights - for example, the right to conduct research and the right not to be researched - based on salient features of the research environment. However, it does not justify an unequal distribution of access to research ethics mechanisms. For example, while Spicker argues that informed consent is not a moral obligation that rests on studying public actions, he holds that researchers are still obligated to respect the general rights of those involved in their research and to ensure appropriate safeguards are in place (Spicker, 2011). What these exact demands are could differ. It seems reasonable to some degree to adjust the ethical demands of research ethics guidelines to the challenges of the research field it aims to guide.

Thirdly, the nature of the research goals between these two categories also does not justify unequal access to research ethics protections and demands. There can be apparent differences between the goals of various forms of research, such as solving social challenges, developing commercial products, or 'mere' academic curiosity. However, it seems unconvincing that these goals *per default* justify a differential treatment since both currently ethically unregulated and regulated research can be conducted for economic purposes, solving social challenges, or satisfying academic curiosity. It seems more reasonable to argue - in line with current practice – that different public or corporate research goals, aims, or needs can present particular reasons that override existing demands, especially when research protections are generally taken not to be

[5] In this paper, I hold risks to create or increase the likelihood of harm (Maheshwari & Nyholm, 2022). Additionally, I define harm not merely in a physical sense, but as any wrongful setback to or thwarting of an interest, such as the violation of a right (Feinberg, 1984).

[6] That is not to say that over-regulation and under-regulation are not serious issues that can both have ethical implications. Rather, this would not be an argument against harmonizing research ethics protections.

absolute and can be overridden by outweighing considerations.

Fourth, do the different professional roles give meaningful conceptual reasons to treat these two categories as distinct? I am not sure. Different professions are often assigned specific ethical obligations associated with their professional role (Wilson & Hunter, 2010). Currently, we hold scientific and publicly funded researchers to have some ethical obligations toward their research participants. In contrast, other researchers are not assigned or do not recognize these professional role-based obligations. However, the distinction between a researcher in a corporate lab and a university lab is not as straightforward as the different professional roles between researchers, teachers, lawyers, doctors, nurses, engineers, etc. To what extent do the researchers who are or are not regulated have substantively different professional roles? What reasons exist for treating these professional roles as researchers substantially differently regarding ethical demands?

One reason that might justify holding scientists to higher standards would be safeguarding the public's trust in scientific research as a social good (Wilson & Hunter, 2010). However, this is an empirical claim. It is not clear that the public trust in science is affected by the current higher ethical demands in the first place (Wilson & Hunter, 2010). However, assuming there is empirical data to back up a casual connection between research ethics regulation and public trust, does the public's trust in the research conducted by non-scientific researchers need not be safeguarded either? Corporations, for example, also seem to benefit from public trust in their researcher to succeed in the market.

Alternatively, scientific or publicly-funded researchers might derive their professional obligations from conducting research with public financial support. Consequently, they might consider they have a particular responsibility to the public, for example, by helping to improve society and not harm it in the process. However, as I pointed out earlier, when public organizations conduct research, such as policy experiments, they are not necessarily held to research ethics regulations despite drawing from public funds. Additionally, while corporate researchers might not depend on public funds (although sometimes on public investments), they still benefit from society in other ways when conducting real-world research, for example, from people interacting with their system or by using public facilities or infrastructures that public funds maintain.

Additionally, I am unsatisfied with the idea that we should *only* extend particular ethical demands to those professional roles that – through historical coincidence – came to view themselves as a particular profession or vocation with specific responsibilities. I am unsatisfied with this because it would mean we distribute regulatory demands based on the willingness of a particular professional community, placing the highest demands on the most willing. This opens the door to rejecting externally imposed regulation based on the argument that it falls outside their professional role's scope. However, professional roles are not static; they evolve over time. The first developments of modern research ethics in the shape of guidelines outlining the responsibilities of medical scientists were also not readily adopted within the professional community. Equally, researchers in public institutions or corporations might feel that their professional duties do not include the demands of research ethics regulations. Nevertheless, we might still have reason to subject them to particular research ethics demands. Moreover, their professional role and associated obligations might change accordingly over time.

Arguably, this point remains somewhat inconclusive. I am sympathetic to the idea that different professional research contexts might present different overriding reasons for the professional obligations of researchers and that these obligations should not be taken to be absolute and can be overriding in the face of sufficient reason. However, it is unclear to me how an appeal to the professional role and duties of a, for example, corporate researcher vis-à-vis a public researcher could be used to argue that research ethics should not be harmonized in the first place. Further work is necessary, however, to work this out further and, for example, ground the professional obligations of researchers in their activities as researchers first and in identifying their employers second.

This, however, is beyond the scope of this paper. Instead, in the next section, parking the question about whether the professional roles of different researchers bring about sufficient reasons to treat them to different ethics regulatory standards, I argue we have good reason to harmonize such practices for other reasons. Specifically, I will argue that the current situation of placing different ethical demands on different categories of research is problematic because it can be exploited through regulatory avoidance of these ethical demands, which comes at the expense of those whom these demands aim to help protect.

5. The argument against regulatory arbitrage

In this section, I argue that the current differential treatment should be amended since it creates the opportunity for what I will call research ethics regulatory arbitrage. That is, the potential avoidance of research ethics burdens by placing research activities outside the current scope of research ethics regulation,[7] which comes at the expense of those whom these demands aim to help protect.

Modern research often involves collaborations between academic, public, and private entities. As Colona argues, modern scientific research "is conducted by a wide variety of actors, including private entities, ranging from small startups to powerful tech giants, as well as governmental and non-profit organizations" (2023, p.1-2). The Ada Lovelace Institute notes that "an increasing amount of AI research involves multi-site studies and public-private partnerships" (Ada Lovelace Institute, 2022, p.7). Real-world experiments with emerging technology are often conducted in transdisciplinary teams and organized in so-called 'real--world laboratories' or 'living labs,' where academic, public, and private parties collaborate to develop and test digital technologies or policy interventions in real-world settings (Ansell & Bartenberger, 2016, 2017).

The current differential research ethical demands are problematic given the collaborative nature of modern research practice because they create a problem of different research ethics standards and responsibilities within a collaborative research project, which can be abused (Ada Lovelace Institute, 2022, p.7). Research parties can avoid taking responsibility for certain research practices by placing them outside the scope of their responsibilities. Since research ethics responsibilities or regulations often end at institutional funding boundaries, this opens up possibilities for research ethics regulatory arbitrage. Regulatory arbitrage occurs when some entity structures its activity around taking advantage of favorable laws. As Colonna notes, it is generally understood as "an avoidance strategy of regulation that is exercised because of a regulatory *inconsistency*" (2023, p.2, italics my own). Research ethics arbitrage would be a subset of regulatory arbitrage focused on meeting research ethics demands. When one category of research is regulated by different ethical demands compared to another, it might be in the interest of those who do not want to adhere to particular burdens of ethical regulation to place or conduct (part of) their research out of the regulatory scope. Data for which scientific researchers might not get research ethics approval could be collected by industry partners operating outside the scope of institutional review boards. At that point, scientific researchers would be working with 'existing' data, which would not prompt the need for ethical review. This would allow academic researchers or transdisciplinary collaborations to circumvent research ethics regulations (Grimmelman, 2015).

An (in)famous example is the Facebook Emotional Contagion study, in which researchers at Facebook, in collaboration with Cornell University, researched emotional contagion through social networks (Kramer et al., 2014). The study was conducted in the following way. Researchers at Facebook would alter the amount of positive or negative posts on the news feeds of certain users to see whether their subsequent

[7] This section will not discuss why and how this differential treatment arose.

posts were affected by this exposure. The 'emotional contagion' then refers to the question of whether exposure to either positive or negative posts would change the emotional state of the user's post positively or negatively. This manipulation resulted in a large set of data that the Cornell researchers were granted access to by Facebook (Flick, 2016). The study generated substantial controversy. Facebook users were unaware that their feed was manipulated, did not give consent to their participation, and the company did not seek ethical review. However, their studies were aimed at manipulating people's emotions. The Cornell University researchers had sought IRB approval, but since data collection was technically done independently from Cornell before their involvement – and they were essentially working with existing data – the IRB judged that no review was necessary (Flick, 2016). I do not wish to suggest that the Cornell University researchers intended to circumvent IRB approval. However, the exact mechanism can be used intentionally to do just so.

In other words, an uneven field of ethical demands across different research domains leaves space for avoiding these demands. This comes at the expense of those whom these demands aim to help protect. This provides a strong case for the current situation to be amended. Research ethics codes, mechanisms, and regulations can thus benefit current unregulated real-world experimentation with digital technologies by promoting collaboration and harmonization across disciplines and industries. However, there are challenges to this goal. I will discuss these briefly in the next section.

6. Going forward: towards a research ethic of real-world research

In the previous two sections, I have argued that the current situation of different research ethics demands between different categories of research is problematic since (1) there are no clear, meaningful justifying reasons to place different demands on currently regulated and unregulated research and (2) the current situation of unequal demands can be exploited by circumventing ethical demands at the costs of person's protection. However, both arguments can swing in two ways. They can be used to justify reducing the regulatory burden on currently regulated research to bring it closer to unregulated research practice, or they can be used to increase the regulatory burden on currently unregulated research. As Grimmelman notes: "To the extent that the Common Rule reflects a consensus about academic research on social media users, it should extend also to corporate research on social media users, because the ethical argument for regulating the latter is at least as strong as the argument for regulating the former … But if corporate social media experiments do not need to worry about informed consent or ethical oversight, we should be having a conversation about exempting academics, too." (2015, p. 254). While I agree with this, I think a stronger case can be made to bring currently unregulated research more in line with currently regulated research. However, this is outside the scope of this paper (see, for example, Hansson (2011) for an argument in favor of this position). Instead, assuming that harmonizing ethical demands means higher ethical demands for ethically unregulated research domains, this section will briefly discuss several opportunities and challenges to transporting research ethics guidelines, protocols, and regulations.

Polonioli and colleagues have discussed the benefits and challenges of several mechanisms for a soft ethics framework for controlled A/B testing, such as (internal) institutional review boards (IRBs), environmental, social, and corporate governance (ESG), the role of conferences, journals, and editorial guidelines, the use of participant compensation and a complementary list of questions to prompt reflection on ethical A/B testing (Polonioli et al., 2023). They rightfully point out that: "companies should not be left alone in trying to elevate their standards of ethical experimentation. Engineers and developers often involved in experiments are not systematically trained in ethics, may perceive ethical considerations as unnecessary red tape, and need to grapple with unavoidable conflicts of interests due to the close link between business and science. To foster compliance with ethical principles, companies need to be properly educated, governed, and incentivized" (Polonioli et al., 2023, p. 679). However, as they admit, a shortcoming of their approach is that they largely hinge on companies' and institutions' *voluntary* adoption of these recommendations.

Whether self-regulation is sufficient has been called into question. As Grimmelman notes: "The history of privacy protections shows that self-regulation without corresponding regulatory oversight is a cruel joke at the expense of users. Unless we start from a place in which social media research is subject to ethical limits, we will not get there. If companies like Facebook and OkCupid know that they must deal with the Common Rule, they will have every incentive to work constructively to fix its imperfections. If they believe that they fall outside it, they will fight tooth and nail to continue in their unregulated ethical free-fire zone" (2015, 259). Some empirical evidence supports this. A study by Stahl et al. which focused on the adoption of Responsible Research and Innovation (RRI) practices in the European ICT industry, concluded that: "innovators recognise some of the ethical and societal concerns associated with their [research and development] activities but their approach is often piecemeal; primary focus is upon the most immediate issues and on legal compliance, to the detriment of broader societal issues and wider challenges" (Stahl et al., 2019). Alternatively, while most researchers can grasp the implications of their research since their behavior is being evaluated, there is an incentive to evaluate it over-favorably (Wilkinson, 2010). This paints a compelling argument to tie research ethical obligations to external audits or legal compliance.

Here, inspiration can be drawn from the EU AI Act and regulatory sandboxes (Buocz et al., 2023; Ranchordas, 2021). The AI Act is a piece of European legislation seeking to regulate the development and deployment of artificial intelligence systems within the European Union. Based on a pre-established risk profile of an AI system linked to its intended purpose, the AI Act makes market approval conditional to specific requirements. However, wishing not to stifle research and innovation, the European AI Act offers exemptions to its regulation for AI research (Colonna, 2023). Given that certain conditions are met, real-world testing with high-risk AI within and outside EU-sanctioned regulatory sandboxes is allowed. Madiega and Van De Pol define regulatory sandboxes as "regulatory tools allowing businesses to test and experiment with new and innovative products, services or business under the supervision of a regulator for a limited period of time" (Madiega & Van de Pol, 2022, p.2). According to the Council, this supervision is done by: "a competent authority which offers providers or prospective providers of AI systems the possibility to develop, train, validate and test, where appropriate in real-world conditions, an innovative AI system, pursuant to a sandbox plan for a limited time under regulatory supervision" (bg). This same 'competent authority' "shall provide, as appropriate, guidance, supervision and support within the sandbox with a view to identifying risks, in particular to fundamental rights, health and safety, testing, mitigation measures, and their effectiveness in relation to the obligations and requirements of this Regulation and, where relevant, other Union and Member States legislation supervised within the sandbox" (1e).

Real-world experimentation outside the regulatory sandbox is also possible, given that certain conditions are met. These include (1) "request informed consent of natural persons to participate in testing in real world conditions", (2) enable oversight by competent authorities, (3) submit a "real-world testing plan .. to competent market surveillance authority", (4) "register the testing in dedicated sections in the EU-wide database", (5) "set limitations on the period for which the testing can be done", (6) "require additional safeguards for persons belonging to certain vulnerable groups", (7) provide "a written agreement defining the roles and responsibilities of prospective providers and deployers and effective oversight by competent personnel involved in the real world testing", (8) "ensure that the predictions, recommendations or decisions of the AI system can be effectively reversed and disregarded", (9) ensure

"that personal data is protected and is deleted when the subjects have withdrawn their consent to participate in the testing without prejudice to their rights as data subjects under the EU data protection law" (72b). Placing similar demands on ethically unregulated research could be a step towards harmonizing research ethics demands.

The benefit of these conditions is that they are necessary for market entry. Companies, governments, and research institutions thus have a financial and legal incentive to abide by them if they want to develop and, eventually, deploy, in this example, a particular AI system. Governments or oversight authorities could use a similar structure to regulate under what conditions researchers could access public space for real-world experimentation. However, they are also quite substantial requirements. They place significant demands on oversight authorities to enact them and researchers to abide by them, thus raising questions of proportionality to regulate various forms of real-world experimentation. Additionally, these conditions would need further conceptual specification (what exactly constitutes 'additional safeguards for vulnerable groups'?), sensitivity to undermining circumstances (what is the value of consent when a participant has limited alternative options available?) and overriding reasons (under what conditions should these demands be allowed to be overridden?).

7. Conclusion

In this paper, I have drawn attention to inconsistencies in the current ethical regulation between various categories of real-world research. Testing technologies under real-world conditions is widespread, yet the ethical issues it raises are often neglected. I have argued that these inconsistencies in ethical demands and protections are problematic. No apparent meaningful difference warrants this differential treatment, and it creates the possibility of regulatory evasion at the cost of public protection. I have grounded my argument in several larger scholarly debates on the limits of current research ethics guidelines and protocols in the face of novel technologies and research formats. I contribute to these debates by drawing attention to a new area needing research ethics reforms. I have briefly discussed several ways forward by drawing inspiration from the AI Act's current regulation on real-world testing of high-risk AI, yet pointed out that these approaches bring about new problems of their own.

CRediT authorship contribution statement

Joost Mollen: Writing – review & editing, Writing – original draft, Investigation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

No funds, grants, or other support were received. The PhD research project of the author is made possible by The Province of South- Holland, The Netherlands. The author has no relevant financial or non-financial interests to disclose. I want to express my thanks to Michael Klenk and the anonymous reviewers for their feedback to help me improve an earlier version of this paper.

References

Ada Lovelace Institute. (2022). Looking before we leap: Ethical review processes for AI and data science research. Available at: https://www.adalovelaceinstitute.org/report/looking-before-we-leap/.

Amnesty International. (2020). *Netherlands: We sense trouble: Automated discrimination and mass surveillance in predictive policing in the Netherlands.* Retrieved March 2022, from https://www.amnesty.org/en/documents/eur35/2971/2020/en.

Ansell, C. K., & Bartenberger, M. (2016). Varieties of experimentalism. *Ecological Economics, 130*, 64–73.

Ansell, C., & Bartenberger, M. (2017). The diversity of experimentation in the experimenting society. *New Perspectives on Technology in Society* (pp. 36–58). Routledge.

Benbunan-Fich, R. (2017). The ethics of online research with unsuspecting users: From A/B testing to C/D experimentation. *Research Ethics, 13*(3-4), 200–218.

Bean, S. (2010). Beyond research exceptionalism: A call for process redesign. *The American Journal of Bioethics, 10*(8), 58–60.

Beerbohm, E., Davis, R., & Kern, A. (2020). The democratic limits of political experiments. *Politics, Philosophy & Economics, 19*(4), 321–342.

Buocz, T., Pfotenhauer, S., & Eisenberger, I. (2023). Regulatory sandboxes in the AI Act: Reconciling innovation and safety? *Law, Innovation and Technology, 15*(2), 357–389.

Calo, R. (2013). Consumer subject review boards: A thought experiment. *Stan. L. Rev. Online, 66*, 97.

Colonna, L. (2023). The AI Act's research exemption: A mechanism for regulatory arbitrage?.

Dainow, B., & Brey, P. (2021). Ethics by design and ethics of use approaches for artificial intelligence. *European Commission DG Research & Innovation RTD*.

DeArman, A. (2019). The wild, wild west: A case study of self-driving vehicle testing in Arizona. *Ariz. L. Rev., 61*, 983.

Desposato, S. (2018). Subjects and scholars' views on the ethics of political science field experiments. *Perspectives on Politics, 16*(3), 739–750.

Feinberg, J. (1984). *Harm to others* (Vol. 1). USA: Oxford University Press.

Flick, C. (2016). Informed consent and the Facebook emotional manipulation study. *Research Ethics, 12*(1), 14–28.

Fried, B. J., Lagunes, P., & Venkataramani, A. (2010). Corruption and inequality at the crossroad: A multimethod study of bribery and discrimination in Latin America. *Latin American Research Review, 45*(1), 76–97.

Galič, M. (2019). Surveillance, privacy and public space in the Stratumseind Living Lab: The smart city debate, beyond data. *Ars Aequi, special issue July/August*.

Grimmelmann, J. (2015). The law and ethics of experiments on social media users. *Colorado Technology Law Journal, 13*, 219.

Hansson, S. O. (2011). Do we need a special ethics for research? *Science and Engineering Ethics, 17*, 21–29.

Humphreys, M. (2015). Reflections on the ethics of social experimentation. *Journal of Globalization and Development, 6*(1), 87–112.

Jouhki, J., Lauk, E., Penttinen, M., Sormanen, N., & Uskali, T. (2016). Facebook's emotional contagion experiment as a challenge to research ethics. *Media and Communication, 4*(4).

Kramer, A., Guillory, J., & Hancock, J. (2014). Experimental evidence of massive-scale emotional contagion through social networks. *Proceedings of the National Academy of Sciences, 111*(24), 8788–8790. https://doi.org/10.1073/pnas.1320040111

MacKay, D. (2018). The ethics of public policy RCTs: The principle of policy equipoise. *Bioethics, 32*(11), 59–67. https://doi.org/10.1111/bioe.12403

Madiega, T., & Van de Pol, A. L.. Artificial intelligence act and regulatory sandboxes. https://www.europarl.europa.eu/RegData/etudes/BRIE/2022/733544/EPRS_BRI(2022)733544_EN.pdf.

Maheshwari, K., & Nyholm, S. (2022). Dominating Risk Impositions. *The Journal of Ethics, 26*(4), 613–637.

McDermott, R., & Hatemi, P. K. (2020). Ethics in field experimentation: A call to establish new standards to protect the public from unwanted manipulation and real harms. *Proceedings of the National Academy of Sciences, 117*(48), 30014–30021.

Miller, F. G. (2010). Striking the right balance in research ethics and regulation. *The American Journal of Bioethics, 10*(8), 65-65.

Moffatt, B. (2010). Not all human subjects research is exceptional. *The American Journal of Bioethics, 10*(8), 62–63.

Mollen, J. (2023). Moving out of the Human Vivarium: Live-in Laboratories and the Right to Withdraw. *Journal of Ethics and Emerging Technologies, 33*(1), 1–22.

Molnar, P. (2020). Technological testing grounds: Migration management experiments and reflections from the ground up. *EDRI, November Available at edri. org/wp-content/uploads/2020/11/Technological-Testing-Grounds. pdf (accessed 20 April 2021)*.

Moreno, M. A., Goniu, N., Moreno, P. S., & Diekema, D. (2013). Ethics of social media research: Common concerns and practical considerations. *Cyberpsychology, Behavior, and Social Networking, 16*(9), 708–713.

Metcalf, J., & Crawford, K. (2016). Where are human subjects in big data research? The emerging ethics divide. *Big Data & Society, 3*(1), Article 2053951716650211.

Pfotenhauer, S., Laurent, B., Papageorgiou, K., & Stilgoe, A. J. (2022). The politics of scaling. *Social Studies of Science, 52*(1), 3–34.

Phillips, T. (2021). Ethics of field experiments. *Annual Review of Political Science, 24*, 277–300.

Polonioli, A., Ghioni, R., Greco, C., Juneja, P., Tagliabue, J., Watson, D., & Floridi, L. (2023). The ethics of online controlled experiments (A/B Testing). *Minds and Machines*, 1–27.

Rahman, H. A. (2024). Unethical online experiments risk real world harm. *Financial Times*. https://www.ft.com/content/de6ab765-35e9-4146-8527-e82795062173.

Rahman, H. A., Weiss, T., & Karunakaran, A. (2023). The experimental hand: How platform-based experimentation reconfigures worker autonomy. *Academy of Management Journal, 66*(6), 1803–1830.

Ranchordas, S. (2021). Experimental regulations for AI: Sandboxes for morals and mores. *University of Groningen Faculty of Law Research Paper*, (7).

Raymond, N. (2019). Safeguards for human studies can't cope with big data. *Nature, 568*(7752). https://doi.org/10.1038/d41586-019-01164-z, 277–277.

Resseguier, A., & Ufert, F. (2023). AI research ethics is in its infancy: the EU's AI Act can make it a grown-up. *Research Ethics*, Article 17470161231220946.
Smuha, N. A. (2021). Beyond the individual: governing AI's societal harm. *Internet Policy Review, 10*(3).
Spicker, P. (2011). Ethical covert research. *Sociology, 45*(1), 118–133. https://doi.org/10.1177/0038038510387195
Stahl, B. C., Borsella, E., Porcari, A., & Mantovani, E. (2019). Responsible innovation in ICT: Challenges for industry. *International Handbook on Responsible Innovation* (pp. 367–378). Edward Elgar Publishing.
Stilgoe, J. (2020). *Who's driving innovation. New Technologies and the Collaborative State*. Cham, Switzerland: Palgrave Macmillan.
Susser, D. (2021). Predictive policing and the ethics of pre-emption. *The Ethics of Policing: New Perspectives on Law Enforcement*, 268–292.
Svensson, S., & Hansson, S. O. (2007). Protecting people in research: a comparison between biomedical and traffic research. *Science and engineering ethics, 13*, 99–115.
Taylor, L. (2021). Exploitation as innovation: research ethics and the governance of experimentation in the urban living lab. *Regional Studies, 55*(12), 1902–1912.
Teele, D. L. (2014). Reflections on the ethics of field experiments. *Field experiments and their critics: essays on the uses and abuses of experimentation in the social sciences* (pp. 115–140).
Vitak, J., Shilton, K., & Ashktorab, Z. (2016). Beyond the Belmont principles: Ethical challenges, practices, and beliefs in the online data research community. In *Proceedings of the 19th ACM conference on computer-supported cooperative work & social computing* (pp. 941–953).
Whitfield, G. (2019). TRENDS: Toward a separate ethics of political field experiments. *Political Research Quarterly, 72*(3), 527–538.
Wilkinson, T. (2010). Assessing the Case for the Regulation of Research. *The American Journal of Bioethics, 10*(8), 63–65.
Wilson, J., & Hunter, D. (2010). Research exceptionalism. *The American Journal of Bioethics, 10*(8), 45–54.
Wood, M. (2020). *OKCupid plays with love in user experiments*. NY Times, 28 July 2014. < https://www.nytimes.com/2014/07/29/technology/okcupid-publishes-findings-of-user-experiments.html. >. Accessed 6 July 2022.
Zimmermann, V. (2023). Smart cities as a testbed for experimenting with humans?-Applying psychological ethical guidelines to smart city interventions. *Ethics and Information Technology, 25*(4), 54.