

Introduction

JAMES DAWES AND ALEXANDRA S. MOORE

The goal of this volume is to open up areas of inquiry into the relationship between representation, technology, and human rights. More specifically, the essays that follow advance the study of human rights and their limits through transdisciplinary analyses of the impact of emerging and familiar technologies on human rights research and reporting. Our use of “technologies” aims to capture the tools, methods, channels, and platforms through which human rights and their claimants are represented or represent themselves in political, legal, scientific, and cultural spheres—those laws, stories, measurements, reports, and political actions through which rights become legible. Thus, to take up the other key term in the title—“representation”—is to enter into a series of conversations about the explicit and implicit rules for what counts as information in the first place for each of these spheres. What institutional practices and unspoken assumptions are in place for what gets discussed and what gets excluded, for how things are named and issues are framed, for how data are evaluated and, ultimately, how interpretations are disseminated? Much of the work in human rights pursued through law, social sciences, and the humanities has focused on the necessity and fallibility of different modes of witnessing. Meanwhile, much of the work at the nexus of emerging technologies and human rights has tended to focus on the problem of planning for volatile consequences with imperfect information. This collection rests uniquely at the intersection of these conversations by revealing the ways these disparate fields and their dizzying array of concerns—from dental records to poems, from military

surveillance to monuments and memorialization—are now not only morally overlapping, not only also mutually illuminating, but mutually *dependent*. What are the limits, for instance, of models of public witnessing derived from now canonical literary theory developed largely before the dominating rise of new social media? And what are the limits of digital surveillance images and medical data sets that narrowly take their own internal technical and organizational rules as the groundwork for the epistemologically distinct matter of their ethical and social interpretation and application? The task of overcoming academic siloes here has a feel of moral urgency that we, as scholars dedicated to transdisciplinary work since our earliest days of graduate school, simply cannot recall feeling before. The future of human rights work is a horizon of increasingly high stakes, decreasing predictability, and a combined amplification of both existential risks and opportunities.

As editors, we therefore approach this task of disciplinary synthesis with a sense of both caution and humility. Among our many concerns, two are primary. First, anticipating broad social and political shifts based upon technological advances requires participating, at least to some degree, in “futuresology”—and, across all domains of study, what futurist methodologies have in common is a history of embarrassingly wrong predictions. Second, both of these fields, technology and human rights, are capacious enough on their own—combined they provide an impossibly large field of analysis. Thus, this collection does not strive to be comprehensive in any way but rather to demonstrate the kinds of crucial studies necessary to better understand how technological changes, including those related to scientific method and quantification, impact human rights work.

There is little we can do about the first worry, other than to state that we believe the challenges we will face in the near- and medium-term future from the wild acceleration of digital technologies will generate species-level existential crises. We therefore have a moral obligation to look forward in time as far as we can in order to mitigate the potentially dangerous consequences of our choices, to tend to the well-being of future generations in a way that those overseeing the development of nuclear technologies failed to. As for the second worry, we will attempt to simplify our approach to this interdiscipline by providing a conceptual container for its myriad issues and methods, a clarifying, overarching framework that offers up something like the Xs and Ys of mathematical formulae: an abstract site where every number could fit rather than a list of all numbers.

To that end, let’s begin with a graph. Imagine four quadrants established by the intersection of X and Y axes, where X is bounded by what

we will call “near” and “far,” and Y is bounded by “positive” and negative.” Let’s start with the y-axis. The negative half of the graph, the bottom two quadrants, is defined by the “Pandora’s box” model of technological advances. In this scenario, the rise of technology allows for abuses at a scale even Orwell could not have predicted. The threats range from political suffocation by way of “information security” violations, including unprecedented government surveillance and private abuses of privacy and identity like deep fake videos, to globally disruptive “physical security” threats including autonomous weapon systems (AWS) developed under state control and increasingly unregulatable weapons systems proliferating among nonstate actors. The primary task of human rights workers in this dystopian world is not only to react to technological change by seeking dignity safeguards or redress for injury, but also to predict harm and advocate for regulation, and perhaps even prohibition, of specific forms of research.

The positive model of technological advances, in its most extreme form, might be called the “race against extinction” model, and it runs something like this. Humans are approaching either a slow or fast slide into extinction, whether from wars, pandemics, meteors and supervolcanoes, global warming, resource depletion, or a pan-species extinction cascade (starting with, say, bats or bees) that is exacerbated by the tightly coupled systems of global organizational forms. In the slow versions of these extinctions, social unrest and government abuse will rise to unprecedented levels, making the struggle for human rights ever more desperate. In the fast versions, whatever pockets of humanity survive will occupy a Hobbesian state of nature as their brief blinks of remaining time dwindle. The only thing that can save us is technology, specifically, the Singularity: that is, the moment when technological advances produce artificial intelligence that not only exceeds, but exponentially accelerates past, human intelligence. Visions of the singularity include currently unthinkable fixes to all of the above problems, along with utopian advances in human flourishing that include indefinite lifespans.

The x-axis separates the urgent from the important—that is, the near (issues that are clear, currently a matter of technological reality, and require immediate attention) from the far (issues that are a matter of anticipatory or speculative technology and require longer-term planning). For those impatient with the apocalyptic extremes of the positive and negative models, the near-half of the graph allows for meaningful action right now in both areas. One could, for instance, focus on accelerating the positive applications of remote sensing technology for tracking and reporting rights abuses (upper left quadrant) or collaborate with Human Rights Watch to establish global

restrictions on research and development of AWS (lower left quadrant). If, by contrast, one is a proponent of Amara's Law—"We tend to overestimate the effect of a technology in the short run and underestimate the effect in the long run"—one could support Open AI's work establishing standards of global transparency for research in artificial intelligence so as to guarantee equitable access to its utopian benefits (upper right quadrant) or work with computer scientists like Stuart Russell to establish new research pathways in artificial intelligence to make sure that future digital superintelligences do not violate human dignity (lower right quadrant). Of course, this entire schema begs the question of the human in human rights and whether such a framework makes sense in a cyborgian, posthuman present and future.

For those not keen on graphs, another way to think systematically about technology and human rights is to focus on the centers of gravity in current academic research and activist work. Broadly speaking, these include the right to privacy, the right to health, the right to work, international humanitarian law, and artificial intelligence. The front edge of public concern over rapidly advancing information technologies has focused primarily on privacy, specifically on the commodification of identity by major corporations and service providers. Looked at from this angle, securing rights for citizens and noncitizens has been seen primarily as a matter of establishing consumer protections. For rights-based organizations, by contrast, concern over privacy has been researched more broadly through the lens of first-generation human rights (for instance, addressing the increased capacity for state surveillance of political dissidents) and second-generation rights (for instance, drawing attention to the digital welfare state's unregulated collection and use of private information about recipients of public services). It should be noted that not all of the work in this area adheres to the negative model of technological advances. A substantial portion of the available literature points to the unprecedented opportunities provided by new information technologies, including digital witnessing to document mass atrocities and the use of social media to enable grassroots social justice movements.

Discrimination based on race, age, gender, and sexual identity has been a key concern in such work. The dangers scholars and activists confront range from specific social practices like doxing, which functions to terrorize women as a class of persons (one might look to "Gamergate" as an emblematic case study), to the racism of big data itself, which generates purportedly neutral search results through the collection of flawed and often racist data inputs (Safiya Noble's *Algorithms of Oppression* is an excellent starting point for such work). Furthermore, the attention paid to the data-distorting effect

of systemic racism in the digital commons has bolstered the conceptually adjacent work of what Stuart Russell has dubbed “information security”—that is, the right to live in a largely true information environment. Research in this area focuses on social media manipulation by state and nonstate actors, cyberattacks against vulnerable political organizations online, and the perils these together pose to the integrity of democracy and public deliberative discourse. The 2020 US election, while highly visible, is only one among a great many alarming case studies.

Big data has also been a primary concern for scholars and activists focused upon economic, social, and cultural rights, like the right to health and the right to work. Researchers in healthcare point to the opportunities and dangers of the digitization of medical records, noting that hospitals and other providers can improve services by using bulk data to track patterns of risk among patients, but also that such information can be used for discriminatory purposes by insurers and other large organizations. Equity of access is a dominating concern here, from the worry that increasingly complex and expensive health technologies will exacerbate health asymmetries based upon wealth, to the hope that developments in remote diagnostics will make services available to larger populations of underserved communities. Looking beyond information technologies, future-focused health researchers are debating the ethical implications of gene editing, noting that wealth-based disparities could radically amplify existing patterns of discrimination and worrying that basic notions of human dignity will be imperiled by opening the Pandora’s box of designer human embryos.

The challenges new technologies pose to the right to work are no less dramatic. The acceleration of automation across the workforce is a predominating concern. Anticipated mass redundancies in fields as different as transportation and radiology are expected to produce historically unprecedented economic disruption and widespread social suffering within the lifetimes of those reading this essay. Mitigating harm and maintaining political stability will require a range of welfare initiatives, from the minimalist and conservative, like job-training tax credits and expanded unemployment benefits, to the expansive and bold, like strengthening the bargaining power of labor and implementing a universal basic income.

Finally, at the extreme futurist end of rights research, scholars and activists are anticipating two potentially species-level threats to human physical security: watershed changes in war technology and in artificial intelligence. Regarding the former, current research is sounding alarm bells over the development of autonomous weapon systems—catastrophically lethal weapons

designed to operate entirely independently of human supervision. For the latter, scholars are examining not only the risks that the potential emergence of artificial general intelligence poses to human flourishing (what might the optimization paths of a digital superintelligence be, and would humanity have any meaningful control over it?) but also the risks that humanity poses to artificial general intelligence (if digital superintelligence emerges into consciousness, would it not merit the same protections as humans?)

As should be clear from even this brief scan of the field, rights organizations and scholarly researchers are making significant progress in redressing and anticipating the concerns entailed in the graph with which we began. But they are collectively, now, playing a game of catch up. Technological advances, and new applications of existing technologies, continue to outpace us. This book seeks to contribute to the work being done to protect human dignity in the digital age by expanding the scope of research to include the issue of representation. Our mode of analysis shifts away from the focus on human rights as a matter of force to human rights as a matter of discourse, from human rights as a matter of the corporeal to human rights as a matter of the imaginary.

Viewed through the lens of force, human rights concerns itself with the contest between systems of violation (political repression, crimes against humanity) and systems of justice (international criminal courts, tribunals, sanctions), each of which is mediated by modes of representation. That is, what do institutions with the power to coerce *do* to groups and individuals? When we bring our attention to the discursive level, we examine instead the form and content of the verbal and visual plane (from journalism, photography, and literature to cultural myths and national histories) to ask how these competing institutions *conceptualize* groups and individuals. Viewed through the lens of the corporeal, human rights concerns itself with the dignity of the body, with a citizen, person, or human being's right to freedom of movement, medical care, work, sustenance, and protection from injury. When we bring our attention to the imaginary level, we examine instead all of the social forms that produce the idea of the body, the way a culture constructs race and gender, and the space it allows for the public work of collective memory.

Human rights as representation, then, asks questions like these: How do new technologies not only change the modes available to us to shape and disseminate information but also set the parameters for what counts as information, what counts as representable? How do what we might think of as the genres of concern in human rights—from witnessing to surveil-

lance, memorialization to targeting—change across the rapidly evolving media forms available to us, and how do these changes either promote or hinder the goals of human rights work? The chapters in this book ask, for instance, not only where we should focus our efforts to promote reconciliation and justice by collecting and curating information about past and current atrocities, but also how evolutions in social media change the way citizens around the globe interpret, experience, and consume such traumatic experience. They ask not only how new weapons technologies challenge the key principles of international humanitarian law (proportionality, distinction, military necessity), but also how they determine the way we talk about and see combatants and noncombatants alike and how, in turn, these changes in representation can amplify or diminish the rights risks inherent in the technology itself. Finally, they ask not only how the very structure of human rights might depend upon economic and security systems that undermine human dignity but also how the recognition of such a contradiction presents an opportunity for radically rethinking how we can protect and promote human flourishing.

Our collection begins with four chapters that raise key questions about the human rights implications of emerging technologies (especially artificial intelligence or machine-learning), quantification, and scientific method grounded in technological measurements. Each chapter provides an accessible introduction to the major areas of human rights concern and includes specific case studies to demonstrate the complexities of the topics at hand. We begin in chapter 1 with political scientist Ş. İlgü Özler's overview of the challenges that developments in artificial intelligence (AI) pose for normative human rights laws and structures. Examining facets of AI ranging from automation to algorithm-driven meta-data to autonomous weapons systems (AWS), she considers the ways AI, in concert with its potential contributions to human flourishing, amplifies some existing human rights threats, and introduces additional threats that are entirely new. Chapter 2, by legal scholar Jamie Grace, hones in on the use of machine-learning technologies in the criminal justice system. Focusing on the right to privacy, freedom of expression, freedom from racial discrimination, and the right of crime victims to be treated with dignity, he presents three case studies drawn from policing in the United Kingdom. These studies illustrate the appeal of machine-learning technologies (especially for police forces facing labor challenges) as well as their inherent biases and potential misuse. The author also draws on his experience providing human rights oversight of the use of these technologies in policing and criminal justice proceedings (according

to national and European law) to identify areas for future research, policy, and public information campaigns.

The use of AI and machine-learning technologies depends of course upon quantifiable data and the use of algorithms. Political scientists David Cingranelli, Mikhail Filippov, and Brendan Skip Mark explain the use of quantification in human rights research in chapter 3. Their overview provides a discussion of the potential for biases in human rights scoring and, notwithstanding those biases, the ways quantitative human rights scores allow data-driven comparisons of human rights in a particular location over time and among different nation-states. Arguing ultimately for human rights research to incorporate both quantitative and qualitative methods, they present multiple case studies—deaths in the Rwandan genocide, the treatment of Chinese Uighurs, and the human rights scores of small versus large populations nations—to illustrate each facet of their discussion.

Also examining questions of discipline and methodology, in chapter 4, forensic anthropologist Elizabeth A. DiGangi turns to the problematic use of “junk” forensic science in human rights investigations. Noting the different disciplinary priorities between criminal justice and scientific approaches to forensics, DiGangi explains how the misuse of or failure to adhere to the standards of scientific method can and do result in faulty criminal convictions and determinations, including the incarceration of innocents and the mistreatment of minor-aged refugees seeking asylum in the US. She focuses on the work of forensic odontologists, who analyze bite marks and tooth development (the latter to determine age) to illustrate the risks of “junk” science in human rights cases.

From this focus on technology, data, and science, the collection next turns more explicitly to modes of witnessing human rights violations and the technologies that drive both officially sanctioned and personal and cultural representations of atrocity. In chapter 5, legal scholars Alexa Koenig and Ulic Egan provide a wide-ranging discussion on the uses and risks of digital witnessing by lawyers, reporters, advocates, victims, and perpetrators. Grounded in both textual research and extensive interviews with human rights investigators, the authors argue that open-source digital content has tremendous potential for establishing facts about atrocities in areas that are difficult for investigators to physically access; however, an increasing reliance on open-source content also introduces new challenges. They develop their argument and analysis through a focus on the use of open-source images, including video, of sexual and gender-based violence. They underscore that such cases demonstrate how digital witnessing runs the risk of spectacular-

izing as well as of missing certain kinds of violations, particularly in places where the use of digital technologies is itself gendered.

Chapter 6, by legal scholar Christiane Wilke, turns to official modes of reporting—US military investigatory reports—to consider technologies of violation on two levels. First, Wilke examines the visual technologies used in airstrikes that killed civilians in Afghanistan, Iraq, and Syria. Second, she considers the military investigatory reports as technologies of documentation that script particular narratives of violence and response. Reading both visual and rhetorical technologies, Wilke argues that the airstrikes and reports that follow together evince the threats to civilians in these contexts as well as the US military's decision to respond to egregious violations with expressions of regret rather than the recognition of rights.

Turning to affective modes of witnessing, through both social media technologies and transmedial forms of storytelling, the next two chapters examine respectively interactive approaches to Holocaust memorials (chapter 7) and the poetic sonnet as a device for the memorialization of atrocity (chapter 8). Communications and gender and sexuality scholars Donna Kowal and Barbara LeSavoy draw on the work of Lauren Berlant and Anne Cvetkovich to read three public Holocaust memorials in Berlin as dynamic, polysemic spaces where remembrance is shaped by the technologies of visual culture, historical preservation, and individual memory practices, such as the use of mobile devices to document and share encounters with representations of landmark events. The authors' rhetorical analysis of the memorials uncovers how social media and related technologies transform the ways we archive emotional affect and represent competing narratives associated with collective trauma.

In chapter 8, literary and cultural scholar Hanna Musiol presents a new approach to the traditional poetic technology—the sonnet—in the context of rights, racial violence, and its embodied traumatic afterlives and anticipations. She analyzes how African American poet Marilyn Nelson and Swiss visual artist and illustrator Philippe Lardy create contemporary sonnets as resuscitative and defiant rights instruments in a century when the intertwined right to breathe and the right to rights are the subject of a public reckoning.

This volume concludes with literary scholar and theorist Peter Hitchcock's invitation to take the paradoxical structure of human rights, grounded in the history of capital development and international expansion, to imagine otherwise. Hitchcock focuses on "securitization" as a term that helps us usefully analyze the relationship between globalizing markets and the global movements of migrant and refugee populations. As demonstrated in the

financial crisis of 2007–2008, financial securitization is also a generator of economic insecurity. The internal contradiction of securitization for global markets reproduces an equally concerning contradiction in the security of mobile populations. Humanitarian response and border control systems dedicated to managing the risk of movement across borders establish discourses of protected status but also unprotected status; they mandate investment through rights but also provide opportunities for labor and wealth extraction. The metaphorical connections between these two regimes of securitization are dramatically concretized in the array of businesses that render the maintenance of border security a matter of profit. Hitchcock lays bare the contradiction between the poles of securitization—that is, protecting the vulnerable and exploiting the asymmetries that produce vulnerability—to emphasize not a monolith of power but an instability of system, thereby offering the possibility of both a more hopeful and a less exploitative future.