The Future of Human Rights

Austin Choi-Fitzpatrick

Carr Center Discussion Paper
The Future of Human Rights

Carr Center for Human Rights Policy
Harvard Kennedy School, Harvard University
September 15, 2022

Austin Choi-Fitzpatrick
Technology and Human Rights Fellow
Carr Center for Human Rights Policy

The views expressed in the Carr Center Discussion Paper Series are those of the author(s) and do not necessarily reflect those of the John F. Kennedy School of Government or of Harvard University. Faculty Research Working Papers have not undergone formal review and approval. Such papers are included in this series to elicit feedback and to encourage debate on important public policy challenges. Copyright belongs to the author(s). These papers may be downloaded for personal use only.
ABSTRACT

Human rights are dynamic, rather than static. The contemporary status quo emerged via a three-phase process, from conceptualization, to clarification, and to consolidation. The present is an interregnum between two significant eras, a fact which the generations metaphor does not adequately capture. The future of human rights will be shaped by how individuals and institutions engage with advances in technologies that transform and extend the mind and body. Particular attention is paid to innovation in superintelligence, social robots, and augmented humans. One implication of this analysis is that changes to the mind and body are likely to transform the subject of rights and to require the development of a more sophisticated rights ecology. Human rights scholars and advocates should engage in a proactive and ambitious program to prepare for such developments. Such efforts will ensure there are rights to clarify and consolidate in the era to come.

KEY WORDS

Human rights, superintelligence, artificial intelligence, social robots, human augmentation, institutions, norms, social change
I. Introduction

This essay takes as its starting point what at first blush appears to be a controversial position: while often described as a distinct field with a singular coherent arc that unfolds over a number of successive generations, the human rights project can instead be thought of as several dense phases of highly contingent activity. This approach is implicitly motivated by a sense that the quest for origin stories matters less than radical breaks with the past, and that explanations focused on contingency help contextualize sequences where they do exist. This approach should provide a more robust framework for thinking about an uncertain future, as discussed in the second half of this essay. The goal is not to critique earlier explanations of origin stories and generational metaphors in particular but to put them into a broader and more forward-looking context. Only then can we begin to develop and sustain the ambitious line of inquiry that an uncertain future demands.

This experimental rethinking of rights’ origins and evolution highlights two implications. First, this argument suggests that efforts to identify particular places or times of origin should be replaced with attention to rich and innovative seasons that made the contemporary status quo possible. Second, it implies that the generations metaphor nests neatly within a broader process-focused approach to rights over time. Overall, the objective is less to tweak scholarly analysis of the past than it is to propose a framework for thinking about the future of rights.

This conversation is critical since, as the old adage has it, past performance does not necessarily predict future results. Thinking about the future will be aided by an exploration of the last time rights were envisioned, debated, tested, and agreed to. To this end, the history of human rights, certainly in the West, should be thought of in terms of a non-teleological process. What exactly this looks like is what this essay is dedicated to exploring.

II. The Origin Story

There is no consensus on when human rights crystallized as a concept in Western thinking, nor whether there are empirically discernable phases in its history. Devin Pendas (2012) suggests that this is due to variation in the methodologies used by historians and scholars—political, legal, or social lenses vary, and different units of analysis provide different platforms from which to use one’s lens. Such debates are illustrative, however, as they suggest four possible origin stories for human rights.

The first explanation is anchored in the notion that human rights are universal with regard to both space and time. This approach, advanced by rights scholars Micheline Ishay (2020) and Paul Gorden Lauren (2013), suggests that the emergence of human rights involves uncovering and discovery. For Lauren, this takes the form of a Whiggish transition from darkness to light, whereas Ishay’s Hegelian approach sees a set of advances, setbacks, syntheses, and slow progress. To the question of where rights came from, then, the answer is that they were already there, either in whole cloth or in need of some stitching.

The second origin story traces rights to the Enlightenment. For example, political philosopher Robert Lamb (2019) argues that the eighteenth century was critical due to the influence of documents like the French Declaration of Rights of Man and the US Declaration of Independence, while Jack Donnelly and Daniel Whelan (2020) suggest that it was the emergent bourgeoisie class of the time that envisioned and fought for a society with universal individual rights.

The third traces rights to the 1940s. For example, historian Elisabeth Borgwardt (2005) suggests that America’s ascension as a superpower in the 1940s alongside the arrival of the Universal Declaration of Human Rights (UDHR) helped create the modern human rights framework (Henkin, 1996).

The fourth origin story traces rights to the 1970s. Notably, historian and legal scholar Samuel Moyn (2012) argues that it was President Jimmy Carter’s inauguration in 1977 that served as the turning point. While Moyn’s work has been well received, some have noted that this argument focuses on the political instruments of human rights—how it is used and by whom and for what reason—rather than the process that made such consolidation possible.

---

2 Some definitions are in order: By dense I mean linked and overlapping. By phase I mean a discernable (and debatable) window of time during which an above-average amount of activity is taking place. By activity I am referring to public-effecting action in individual, collective, and institutional forms.

3 Of course, Hans Peter Schmidt is right to note that human rights discourse has been colonial, white, and paternalistic.

4 I am indebted to Devin Pendas’s (2012) conceptualization of these camps, as it includes the category of “transhistorical rights” (i.e., Lauren and Ishay), which my original analysis had omitted.

5 For an extended discussion of Lauren’s work see Afshari (2007), and for a discussion of Afshari’s work see Whelan (2019).
Rather than inquiring after the origin point of human rights—a worthy effort that this study relies on—my goal here is different. I suggest that different origin stories represent a number of dense phases of activity that, in retrospect, can be thought of as building on one another. The causal argument that underlies this claim is straightforward: innovative individual and institutional ideas and action emerge in response to particular cultural, economic, political, and technological moments in time. What happens to the debate over origins if we focus on continua, rather than dichotomies?

III. The Generations Story

In an effort to historicize human rights, the Czech jurist Karel Vasak (1977) suggested that human rights had passed through three generations. The first generation, Vasak argued in an influential essay, involved the protection of liberty and participation in public life; had its roots in the Magna Carta, English Bill of Rights, French Declaration of the Rights of Man and of the Citizen, and the United States’ Bill of Rights; and was most clearly visible in the Universal Declaration of Human Rights’ Articles 3–21, the International Covenant on Civil and Political Rights, and the European Convention on Human Rights.

The second generation, according to Vasak, focused on economic, social, and cultural rights, and came into its own after World War II. The emphasis here was on citizenry rights and a basket of benefits disbursed by the welfare state (especially employment, food, housing, healthcare, social security, and unemployment). Second generation rights were to be found in the UDHR Articles 22–28, as well as in the International Covenant on Civil and Political Rights, and the European Convention on Human Rights.

The third generation of human rights, Vasak argued, went beyond civil and social rights and involved solidarity, especially related to self-determination, collective rights, economic and social development, cultural heritage, and natural resources. Third generation rights were protected by instruments like the African Charter on Human and Peoples’ Rights, as well as both the Stockholm Declaration and the Rio Declaration on the environment.

The source of these three generations, Vasak suggested, was the French Revolution, and its call for liberté, égalité, and fraternité. The generations metaphor suggests something of parentage and implies something about sequencing: a Revolution as progenitor, the first generation begat the second, and the second generation begat the third.

Advancements in communication technology have led some analysts to propose the adoption of a fourth generation of human rights. This would likely include the right to equal access to computing and digital spaces, as well as the right to digital self-determination and digital security.

This metaphor has proven useful. “Generations” is a familiar heuristic, helpful in unpacking an otherwise complex process. Human rights emerged from a dynamic historical period (the eighteenth century), were beset by cataclysmic global conflicts, were buffeted by competing force fields during the Cold War, and are subject to debates between the “West” and the “Rest.”

This metaphor has largely gone unchallenged, as the variables in the equation—nation-states, major institutions, people and peoples, and liberté, égalité, and fraternité—have not fundamentally changed in composition. Empires have risen and fallen, the number and type of claimants have expanded, and means of enforcement, both hard and soft, have evolved, but the underlying phenomena has proven rather durable. While the generations metaphor has proven useful in explaining the past, a different approach may be needed to help us think about the future.

---

6 It may be helpful to note here that I believe this causal argument applies to each of the three phases that comprised the past era, just as it applies to the moment we are in, which I argue is a liminal stage between eras. I am grateful to Maria Carnovale for pushing me to clarify this point.

7 This simple advice, given in a graduate workshop by Jack Donnelly many moons ago, has informed most of my work.
IV. A Story of Radical Breaks, Temporality, and Process

The rallying cry of the French Revolution represented a radical break with the past. Liberté, égalité, and fraternité as public claims, the National Assembly as an institutional act, and the storming of the Bastille as a public act sent out into the world a new set of ideas about who is deserving of rights and what forms those rights should take (Scott, 2018). Ideas that emanated outward from that center have animated all subsequent human rights efforts. It is easy to see why Vasak placed the French Revolution at the heart of his generational framing. Of course, the debate over origins reminds us that the French Revolution was but one of several possible progenitors. What’s more, the generations metaphor implies an ordering of impact that is harder to support with empirical analysis (it could be argued that liberté and égalité were twins separated at birth, the former adopted by Britain and America and the latter adopted by the Soviet Union).

This essay steps away from both the origins debate and the genealogical metaphor in order to make two arguments. First, rather than debating origins, scholars should focus on the conditions that engender radical breaks with the status quo. Doing so emphasizes the importance of the Enlightenment for Hunt and the Axial Age for Ishay. The second argument is that the generations metaphor nests neatly within a broader process-oriented effort to understand how rights appear over time. My goal is not to make new enemies but to find new friends focused on better understanding our rapidly changing world. Sustained debate, for example, could fruitfully focus on whether or not we are in a transformative moment (which I suspect we are) or whether we are undergoing another stage of growth (in which case the generations metaphor remains apt).

If we are indeed on the cusp of a revolution, then it requires us to adopt more dynamic and contingent stories and metaphors. I do not think one single essay can do much more than start a conversation. My opening gambit, then, is a hypothetical alternative history, focused on contingency and attentive to process. What if we instead imagine the last two hundred years in terms of three phases?

CONCEPTUALIZATION

Starting in the eighteenth century, new ideas in the republic of letters and facts on the ground worked together to produce innovative and emergent theological and political conceptualizations of humanity as well as novel institutional structures in states and markets. Enlightenment conceptualizations of humanity were more emergent than novel, I hasten to add, as thinkers pulled together many ideas that had been around for a very long time but extended them from science and philosophy in order to advance politics (Ishay, 2020). This phase extends, I argue here, into the mid-nineteenth century. Such a periodization leaves room for a range of activities identified by others and briefly mentioned earlier, including social movements (Simmons, 2009; Tilly, 2004), an emergent bourgeoisie class in Europe (Donnelly & Whelan, 2020), the important secularization of theological natural rights theories in the eighteenth century (Lamb, 2019), and the emergence of the epistolary novel, which generalized empathy in ways that laid the groundwork for subsequent arguments about universalization (Hunt, 2007).

CLARIFICATION

By the nineteenth century, new ideas and facts on the ground worked together to clarify these earlier concepts and to begin operationalizing them in more tangible social, political, and economic identities and collectivities. This phase extends, I argue, into the mid-twentieth century. This timeframe leaves room for important transnational social movements, including multiple waves of collective action calling for the abolition of slavery (Martinez, 2011; Sinha, 2016) and for women’s suffrage (Daley & Nolan, 1994; Mayhall, 2003). National projects—including unions, utopian societies, socialist and communist parties and regimes, and declarations of rights (in France, Russia, and the United States)—as well as experimental institutional configurations like the League of Nations represented a wave of innovative activity that served to operationalize and field-test ideas that had emerged in the conceptualization phase. The result was a host of new social, political, and economic collective action movements, radically expanded notions of citizenship, and the beginning of the end for European colonial efforts. 8

CONSOLIDATION

Starting in the mid-twentieth century, new geopolitical realities intersected with a host of emancipatory and collective action experiments from the clarification phase and facilita-

---

8 I am grateful to Everard Meade for pointing out that states, polities, and social movements used some of these same ideas for a wide array of purposes. For example, humanitarianism pulled in a parallel direction with human rights in the near term but offered a fundamentally different theory of the case than human rights, over the long term. Other ideas, central to the origins of human rights, were mobilized to create new forms of slavery, colonialism, and social hierarchy. The resulting process is far more contingent or dialectical than it is predictive or linear.
tated the emergence of innovative norms and institutions to reinforce a consolidated and bureaucratized roster of universal and indivisible rights (Donnelly, 2013; Whelan, 2011). Here we find the aftermath of the Second World War (Glendon, 2001; Henkin, 1996); America’s ascension during the post-war period (Borgwardt, 2007); the drafting of the UDHR laying the groundwork for the modern consolidated human rights framework (Borgwardt, 2005); the general role of American identity and interest in this phase and the particular role that Jimmy Carter played in ensuring human rights’ place in American foreign policy (Moyn, 2012); the slow collapse of the welfare state and increased awareness of the Holocaust (Moyn, 2012); the rise of Amnesty International (Neier, 2012) and Human Rights Watch (Hopgood, 2013) in particular and grassroots advocacy in general (Quataert, 2009); and the emergence of rights as a concept within global politics, notably during the Kosovo War of 1998–1999, which ushered in a new period of humanitarian intervention (Hoffmann, 2016) and efforts to articulate the international community’s responsibility to protect vulnerable individuals (Evans, 2009). Each one of these actions or processes further enmeshed rights rules and language into national and international norms and institutions.9

In sum, initial upheaval in science and technology, on one hand, and ideas and institutions, on the other, led to novel norms at the conceptualization stage, which had an impact on subsequent opportunities and decisions related to clarification—which in turn led, especially after the horrors of the Second World War, to a consolidation of rights ideas into rules, norms (post Second World War), and institutions (in the 1970s, per Moyn). This approach implies neither path dependence nor complete range of motion but instead something closer to Anthony Giddens’s (1984) structuration in that it acknowledges the role that both structure and agents play in creating and reproducing social systems. The implication is that dense phases of activity are not automatic extensions of the moment that came before but emerge from a wrestling with a particular moment in time using the conceptual, institutional, and material tools at hand. Clearly, ideas, norms, and institutions were evolving across each phase, before consolidating into what I am here calling an “era.”10

We are now in a position to reflect on the generations approach. The model has done solid work helping a generation of thinkers to organize a welter of activity into a familiar metaphor. My sense is that the French Revolution (like the Axial Age before it) repre-

“Human rights emerge at the intersection of precedent and opportunity.”

sented a radical break with the past and that it took time for ideas to spread, be clarified, and to then be consolidated into norms and institutions. Clarifying debates—especially between the United States and the Soviet Union—gave us competing theories of the case, notably conceptualized as first and second generation rights. This is anachronistic, however, as those first and second generations competed with one another on the world stage, suggesting that they might more accurately be conceptualized (to continue the family metaphor) as siblings.

This reconceptualization of origins and evolutions may be important for how we think about, plan for, and actually build what comes next. Transformations in the geopolitical, energy, ecological, and biotechnological spheres will almost certainly transform the human rights status quo and will lead to decisions at the conceptualization stage that have knock-on effects at the clarification and consolidation stages of rights evolution.

Central to this effort is the argument that human rights is a dynamic field, rather than a static entity, and this dynamism is less teleological than the generators metaphor implies. The implication is that we are perhaps more likely to look for the next generation than to anticipate the ways that rights are likely to be buffered by external forces and factors—to ebb, flow, evolve, or even to disappear. Further, I argue that we should develop sensitivity to and curiosity about how and when change happens and what implications this change holds for human rights theory and practice.

I have offered an illustrative three-phase process in order to advance the argument that human rights history is composed of moments of innovation in response to changing conditions on the ground. The conceptualization phase saw innovation around the individual, clarification saw the innovative enumeration of rights, and consolidation saw the emergence of new institutions focused on upholding and enforcing rights.

The primary implication is the most obvious: the world is not static. A secondary implication is that human rights emerge at the intersection of precedent and opportunity, and in response to

9 Steven Hopgood can be imagined arguing that this consolidation represents the shift from grassroots human rights to institutional Human Rights, whereas Kathryn Sikkink can be imagined arguing that this institutionalization is the moment that rights have to be taken seriously within the international system. This essay suggests these arguments are two sides of the same coin, as human rights has become an important coin of the realm.

10 This point is crucial to emphasize, lest I be misunderstood to be arguing that somehow the “conceptualization” stage was the one with ideas, whereas the “consolidation” stage merely created institutions out of what ideas there were to be had. It is more accurate to imagine that ideas are constantly emerging and evolving, but that not all survive—indeed, all possible ideas that have been mooted cannot survive if norms are to emerge. And from there not every norm is instantiable in institutional form—Unger (2015) was right, after all, that humans must choose among worlds.
broader contexts and conditions. It is true that, across each of these phases, war and colonial expansion accelerated the dispersion of people and the diffusion of ideas, laying the groundwork for initial extraction and exploitation and subsequent emancipation movements that turned civilizational arguments on their heads. I focus here on a transformation less attended to by rights scholars: technological change.

Developments in the technologies of the body and mind—especially human enhancement, superintelligence, and social robots—are likely to radically impact much of life as we know it, including individuals, society, state, and economy, and human rights along with them. It may be that these inflection points reify the status quo, with the prime example being the 2008 financial crisis, which further entrenched capital rather than liberating labor (Helleiner, 2014). But maybe not; only time will tell.

What we call human rights represent the conceptual, institutional, and normative working out of challenges that emerge from social, political, economic, and, increasingly, ecological developments (which themselves are interrelated). This “working out” is not one-off but is instead dynamic, evolving, and emergent—a point my conceptualization, clarification, and consolidation heuristic is intended to emphasize.

The earliest days of greatest uncertainty—when significant complex change is on the horizon—represent a liminal and in-between moment during which novel alternative conceptualizations emerge to challenge the increasingly brittle status quo (Van Gennep, 2013). Old ideas are on a precipice, but a host of new realities and possibilities suggest that new ideas are needed if rights are going to evolve in this new era.

We have now laid sufficient groundwork to ask a simple question: What is next for human rights? I am not, of course, the only one to notice that big changes require fresh appraisals (Brysk, 2018). An important debate between Steven Hopgood’s The Endtimes of Human Rights and Kathryn Sikkink’s Evidence for Hope has focused on the implications of changes in international politics, while scholars like Mathias Risse (n.d., 2019, 2021), William Schulz and Sushma Raman (2020), and James Dawes (2020) point to the challenges and opportunities posed by emerging technologies. My goal here is to put all this activity into what I understand to be its larger perspective and in this way lay the groundwork for debate about this hypothesis and its possible implications.

V. Tectonic Shifts

The contemporary human rights regime is the product of innovation across individual ideas, collective action struggles, norms and values, and institutional forms. Rights, as idea and regime, gain relevance at the intersection of novel socio-political and socio-technical developments and emergent normative actions and ideas.

Social movement scholars have focused on the importance of collective action in driving the adoption of various human rights objectives, including citizenship, gender, sexuality, voting, civil, and environmental rights (Tarrow, 2011; Tilly, 1978). This emphasis on contention is important, as agency represents a necessary condition for the emergence of new rights and recognition—indeed, power concedes nothing without a demand.

But this emphasis is also incomplete, as it underspecifies the significance of broader systemic changes. In particular, changes to the means and system of production are overlooked by the more agency-centric accounts offered by most of my colleagues focused on social movements for human rights. Exceptions (Choi-Fitzpatrick, 2020; Della Porta, 2015; Hetland & Goodwin, 2013) prove the rule. These exceptions affirm the importance of the emergent normative actions and ideas mentioned above but direct additional attention to the importance of socio-political and socio-technical developments, including advances in science and technology.

“Old ideas are on a precipice, but a host of new realities and possibilities suggest that new ideas are needed if rights are going to evolve in this new era.”

---

11 I am indebted to Elizabeth A. Bennett for noting this alternate hypothesis. I would note that changes in form are not changes in function—it is possible for significant aspects of the status quo (e.g., capital’s power over capital) to survive under significant upheaval.
Human rights as we understand them today emerged in direct conversation with broad and destabilizing changes in the technological, scientific, and economic spheres. Culture and politics are in a complicated relationship with science and technology (i.e., tightly braided with and emergent from) and are imbricated in and co-constitutive with innovation in these fields. The last major, and most extensively documented, technological innovation event was the product of the military, financial (i.e., capitalist), and cultural revolutions that produced the modern era. That technological revolution played an important role in upending the religious, political, social, economic, and ecological status quo. This upending created new cultural, political, and economic identities, collectivities, entities, and regulatory systems. These tectonic changes made necessary new conversations about the arrangements between employer and worker and between rulers in the early days and later between classes of worker and owners of capital and between the ruling elite and a potentially mass polity, as well as between the center and the periphery.

Such are the three phases—conceptualization, clarification, and consolidation—that brought us to the present moment. Together they define the era of human rights as we understand and debate it. It is my thesis that we are now in a liminal space between two large eras.

The conceptualization, clarification, and consolidation phases of human rights were shaped by globe-spanning events like the “age of discovery,” colonialism, anti-colonialism, nationalism, steam and petroleum economies, and so forth. Humans and their biome were the terrain on which this drama unfolds. This is not the time to parse whether the broader terrain is best thought of using Nikolai Kondratieff’s (1979) long waves in economic life, Simon Smith Kuznets’s (1930) shorter economic waves, Alvin Toffler’s (1980) information-age Third Wave, Klaus Schwab’s (2017) digital Fourth Industrial Revolution, Max Tegmark’s (2017) Life 3.0, or any other overarching theory.

I am more interested in establishing the notion that the present is an interregnum between eras. Put crudely, everything we know about human rights is in fact the product of the conceptualization, clarification, and consolidation processes that emerged in dialogue with this prior dispensation. This is not the time to resurrect a deterministic approach to technology’s role in social change, but we would be wise to explore whether we are in the process of exiting this long and noteworthy era.

Of course, to exit an era is not to enter a new one. If the present is an interregnum, then we can imagine, and perhaps expect, to see a new set of interactions between technological, cultural, and political innovation and new ideas about what this means and how these interactions should play out, and we can perhaps also imagine the emergence of a revolutionary break with the past, which will upend the status quo and create new identities, collectivities, entities, and systems. I want to emphasize the possibility that this new drama could unfold in, through, and beyond humans and their biome. In other words, humans are actors in this process, the subject of this process, and perhaps ultimately a remnant of this process.

What are human rights in this scenario? One is reminded of Antonio Gramsci’s observation that “the old is dying and the new has yet to be born.” What shape the future takes is uncertain, but it is likely to be defined by significant changes. These changes are co-constitutive in nature, combinatorial in impact, and emergent in outcome. What shape the future takes is uncertain, but it is likely to be defined by significant changes. These changes are co-constitutive in nature, combinatorial in impact, and emergent in outcome.

An insightful colleague, upon reading a draft of this argument, asked me, “What is, for human rights, the death rattle? What is the [newborn] baby?” This is a provocative line of inquiry that allows me to clarify two points. First, I am not suggesting that we should prepare for a wake—nobody in the human rights world is seriously arguing that we are in the endtimes of human rights. Rather, there is important talk of backsliding, watering down, doubling down—all are forms of evolution. The second question—how will we recognize the transformative—is exactly the question of the hour. Technology-based changes to the mind and body that raise important questions about whom rights apply to. The rights community should equip itself to answer this question.

---

12 There is a significant debate about the causal direction of this relationship, perhaps best recognized in debates about technological determinism. This project will sidestep any debate over origins in order to explore the complexity of immediate and subsequent consequences, which we believe to be socio-technical in nature.

13 I am indebted to Elizabeth A. Bennett for this insight.
“We would rather imagine... a world perhaps in which ‘human’ as shorthand for Homo sapiens is replaced by a more generous notion of personhood, which evolves to include the many lifeforms capable of flourishing and deserving of respect.”
VI. Technologies of the Mind and Body

This essay is meant to spark debate. It is not meant to present a definitive treatment of human rights’ origin story. Nor is it meant to serve as a map for things to come. Most humans are bad at predicting the future, and academics are worse, bolstered as we are by the confidence and conviction borrowed from knowledge. This does not mean, however, that we cannot begin to sketch the contours of forces likely to determine social, economic, and political life in the decades to come.

A number of important trends—China’s rise, changing energy systems, and increased climate change, among them—rightly receive significant attention. My goal here, however, is to focus attention on developments that I believe are important for rights futures but about which far less has been written: accelerated technological changes to the mind and the body, which are likely to change the subject of rights, notably through the rise of a more complex rights ecosystem.

Human rights pertain to humans. The status quo has revolved around the assumption that the human person is a relatively stable category and the sole unit of analysis. Liberté, égalité, and fraternité—for humans. Yet recent extensions of human rights tools to rivers and animals will soon be joined by debates related to what rights, if any, should be accorded to augmented humans, superintelligence, and social robots.

A definitional detour is in order. By augmented humans I mean a human person who has undergone enhancements such that their mental and/or physical capacities are substantially greater than an unenhanced human person (Buchanan, 2011; Savulescu & Bostrom, 2009). Advances in this area are being pursued along multiple lines, especially transhumanism (More & Vita-More, 2013), life extension (Kurzweil, 2005), and moral enhancement (Persson & Savulescu, 2012), as well as a range of nearer-term medical interventions like brain-computer interfaces (Wolpaw et al., 2000), deep brain stimulation (Perlmuter & Mink, 2006), and retinal implants (Beauchamp et al., 2020; Zrenner, 2002).

The term superintelligence is used to refer to the ability of an intelligent nonhuman agent to understand or learn any intellectual task that a human being can (sometimes also referred to as artificial general intelligence).14 Increased interest in and commentary on artificial intelligence has led to decreased clarity about what the term refers to, with some using the term to refer to superintelligence while others use it to describe machine-learning tools directed at enormous datasets. These popular usages are themselves separate from debates over whether humans will transcend biology (Kurzweil, 2005) or if human consciousness, and freedom and agency along with it, is irreducible to synaptic firing (Tallis, 2016). Important for our purposes is the notion that superintelligence lays the groundwork for a nonhuman agent to demonstrate cognitive abilities and a personality that are indistinguishable from human abilities and personalities. The implications for human rights are only beginning to be explored, notably by (Risse, n.d., 2019, 2021) and (Dawes, 2020).

While the broad term robot refers to a machine capable of carrying out a complex series of actions automatically, my focus here is on social robots, by which I mean robots that take a human form. While social robots needn’t take humanoid form, the possibility that similar (or identical-seeming) corporal entities could inhabit the social world alongside organic humans raises important political, economic, social, and cultural questions. Most of the questions raised by social robots are related to the possibility that a particular physical configuration (i.e., humanoid form) be combined with a particular cognitive capacity (i.e., superintelligence) such that the emergent entity is physically and mentally undistinguishable from what I am inelegantly calling “organic humans.”

Here I should note that neither this definition nor my approach is focused on the near-term dilemmas of automation and technological employment. These are important issues, but both the problem (machines displacing humans) and the proposed solutions (unions and income guarantees) are familiar. The purpose of this article, by contrast, is to draw attention to less familiar challenges in order to inspire vibrant debate and highlight the need for innovative and generative work on solutions.

I am still looking for a term for technological innovation in our minds and our bodies, in particular through the development of augmented humans, superintelligence, and social robots.15 The approach I adopt here intentionally clusters fields that admittedly deserve individual attention, but it does so for good reason: the fields are likely to have an emergent effect on the subject of most human rights interest—the individual.

The status quo revolves around the (at least) two-century-long notion that the proper and exclusive subject of rights are individu-

---


15 Two caveats must immediately be noted. First, enhanced human, superintelligence, and social robot are all concepts held in contrast to a baseline: the human. This baseline is blurry at best. The notion that there is a single benchmark for “human-level intelligence” throws important flags related to age, economic equality, neurotypicality, caloric equality, and so forth. The notion of the “un-enhanced body” is similarly problematic: when was the halcyon era of the body and mind, untouched by ritual scarification, body modification, medicinal intervention, and serotonergic hallucinogens? The dilemma only compounds with the advent of modern scientific approaches to medicine, which includes stem-cell-based treatments, artificial limbs, pacemakers, and augmented reality lenses. A second caveat: this article is being written in the middle of a great welter of activity, as an incredible amount of money is being poured into these three interconnected fields. In the process, many terms go undefined or are used in confusing and overlapping ways. To read a book with “robots” in the title, for example, is often to read an assessment of advances in intelligence.
al humans. Advances in the technologies of mind and body have the potential to significantly disrupt the norms and institutions that support this approach. During the consolidation phase described earlier, activists were hard at work to ensure that people with marginalized gender, sexuality, ethnic, and religious identities were recognized as human, and the eventual extension of human rights to all bipedal Homo sapiens was broadly accepted (by human rights advocates, if not by the general public and decision-makers).

A fresh wave of innovation will require a reconceptualization of the area of philosophy, politics, and policy currently occupied by “human rights,” as breakthrough technologies, capacities, and identities trouble seemingly fixed categories. The implications for human rights are impossible to predict, but several contours are visible on the horizon.

First, these developments suggest that the search for origins is better thought of as the exploration of radical breaks with the past, and sequences of events better thought of processually than generationally. Just when it seems this struggle has led to a consensus, the technological enhancement or replacement of our minds and bodies suggests that there is much more work to be done.

Second, these innovations suggest that the mind and the body might have separate trajectories. Advances in whole brain emulation and superintelligence, on the one hand, and cyborg enhancement and life extension, on the other, will lead to a debate over whether human rights are transhistorical, deeply connected to consciousness as understood by humanity’s earliest philosophical commitments (Ishay, 2020), are rooted in secular trends and political developments (Moyn, 2012), or are tied to our embodiment (Turner, 1993, 2021). This overly simplistic reduction of several longstanding and contentious scholarly conversations is helpful insofar as it highlights some of the fundamental dilemmas that synthetic bodies and minds pose: If my mind is uploaded to a synthetic body, and my discarded body is loaded with a synthetic mind, then which of the entities is human, and which path do one’s rights follow? Nascent scholarship on the promise and peril of superintelligence (Bostrom, 2014) is being joined by a robot ethics, robot rights, and on human companionship with robots (Danaher, 2017; Levy, 2009).

Thirdly, and perhaps most profoundly, these innovations complicate human as the unit of analysis. As rights approaches are being extended to animals and rivers (Schulz & Raman, 2020), new questions are being asked closer to home, as this section demonstrates. New and more generous conceptualizations of the world are needed. Donna Haraway (2016) suggests radically novel approaches to living (“making kin,” in her conceptualization) with others with the goal of multispecies flourishing. This tracks neatly with the more expansive notion of flourishing advanced by Martha Nussbaum (1998) in Cultivating Humanity. There, Nussbaum suggests a number of approaches to multispecies flourishing.

Additional questions abound. What is the relationship between flourishing and consciousness, self-awareness, pain, and embodiment? What of embodied pain if it is experienced by a nonhuman mind?

VII. Discussion

Human rights are created, rather than discovered or begat. Following a radical break with the past, human rights have been forged out of a process that we can imagine in three waves: from early conceptualization, to challenging clarification, to eventual consolidation. There is not telos here, only history, and the question I intend this essay to raise is whether this past process might represent a heuristic for thinking about the current moment as a liminal space or as an interregnum. Doing so raises significant questions about what’s next, but what holds my attention are advances in new technologies of mind and body (augmented humans, superintelligence, and social robots).

It bears noting that this liminal moment is also facilitated by three other great changes in the form of changing global power structures (the rise of China), changing energy regimes (from oil to a broader and more sustainable portfolio), rising temperatures (anthropogenic climate change), and emerging technological advances in the mind and body (augmented humans, superintelligence, and social robots).

Taken individually, these factors pose a significant challenge to the system as we understand it. Taken together, however, these changes represent a crisis for the status quo. The exact form of the crisis is certainly unforeseeable, as anyone who predicted capitalism’s collapse in the aftermath of the 2008 financial crisis can surely tell you. Nevertheless, significant change is on the horizon. In figure 1, I do my best to sketch one of the implications of this argument.

Again, my intent is not to resurrect a deterministic approach to technology’s role in social change. Nor is it to focus on any uniform or linear probabilistic path. Rather, the goal is to flag the possibility that a great undertaking, the human rights enterprise, is at a significant crossroads.

It may help to unpack each of these factors and muse over possible implications. The conceptualization, clarification, consolidation heuristic has been sufficiently described. This illustration prompts the pragmatic possibility that we are alive in the middle of a great field, able to choose paths, rather than the structuralist certainty that we are riding on a train that is fixed to a track. The significant changes described earlier will require an innovative and creative response from the rights community.

If current trends hold, the next two decades are likely to be a fecund and productive time, with many ideas competing for supremacy at the very moment that their underlying assumptions and starting propositions vary wildly or are wholly incompatible. It is at such a moment that the entire enterprise can collapse under the weight of its own contradictions, or a clarification process can begin to tease signal from the noise, identifying new directions for philosophical inquiry, producing clear lines of debate, and posing novel questions for empirical validation.

Augmented Humans—I use the (admittedly overly broad) term augmented humans to describe efforts to increase a human person’s mental and/or physical capacities through radical technologies like those advocated by the transhumanist movement (e.g., cryonics, 3D bioprinting, mind uploading, and telomere lengthening) as well as the pursuit of moral enhancement, brain-computer interfaces, and the like. As these technologies advance, some may start to shade into more mainstream developments from the biotech industry, as recent innovation in prosthetic technology and Elon Musk’s investment in the brain-computer interface firm...
Neuralink make clear. From a human rights standpoint, these efforts have the effect of further accelerating inequality.

Life expectancy is already dramatically distorted by inequality, with the richest 1% of Americans outliving by fifteen years the poorest 1% (Chetty et al., 2016). New (and expensive) innovations in the augmentation space are likely to further accelerate this trend, in the United States and around the world. Recent novels are illustrative. Radically extending a lifespan could have the effect of creating a Methuselah class, which derives incredible benefits from longevity, compounded interest not the least among them. This dilemma is explored at length in Richard Morgan’s cyberpunk novel Altered Carbon. Likewise, radically increasing cognitive capacity could have the effect of creating a Lifted class, which secures outsized benefits from enhanced (but not super-) intelligence, as seen in Kazuo Ishiguro’s novel Klara and the Sun.

Social Robots and Superintelligence—if startups and cyberpunk novels offer one set of puzzles, a cursory skim of the Universal Declaration of Human Rights (UDHR) offers another. Asking how concepts advanced in the UDHR’s thirty articles might interface with advances in social robotics and more-than-human intelligence raises far more questions than answers.

Article 1, for example, reads as follows: “All human beings are born free and equal in dignity and rights. They are endowed with reason and conscience and should act towards one another in a spirit of brotherhood.” Here we find concepts crucial to the human rights enterprise that are very likely to be soon challenged in courts of law (is “human being” a term only describing Homo sapiens?), debated by philosophers (are “dignity,” “reason,” and “conscience” things that only humans have, when not all humans actually have them?, a question well developed by Martha Nussbaum), and operationalized by social attitudes and behavior (what happens if people decide to “act towards [social robots] in a spirit of brotherhood?”). Article 4, for example, reads as follows: “No one shall be held in slavery or servitude; slavery and the slave trade shall be prohibited in all their forms.” It holds my attention, in part, as a result of my research on the contemporary anti-slavery movement.11 On one end of the spectrum, we have scholars writing explicitly in favor of enslaving robots (Bryson, 2010; Chetty et al., 2016); at the other end, we have those who argue that robot prostitutes, far from providing a viable alternative to the human version (Levy, 2009; Levy & Loebner, 2007) are yet another form of exploitation (Richardson, 2016). In between we have those who just aren’t sure (Petersen, 2007), a host of scholars suggesting we should wait to see what exactly is possible. Between these two poles lie experiments like Cybrothel, a Berlin-based operation that offers to customers the opportunity to have sex with Kokeshi, a silicon doll. For an extra fee, the customer can pay to have Alexis Smith, the artist who runs the installation, speak for Kokeshi. Is Alexis a sex worker? Is Kokeshi a victim? Should renting Kokeshi be illegal wherever prostitution is illegal? Answering questions like these will occupy both the anti-trafficking movement as well as organized labor for decades to come.

VIII. Challenges and Opportunities

This essay is more a call to action than it is a philosophical exercise or empirical intervention. The core argument is that technological innovation is likely to require the human rights community to have new conversations, expand its agenda, set new goals, cultivate new interlocutors, develop new tools, expand its resources, and create new avenues for coordination and collective action.

New Conversations—The first and most crucial debate is over whether the rights cannon can be expanded, whether it should be expanded, and, if we’ve answered in the affirmative, how rights might be expanded. Operationalized for the rights of social robots and superintelligence, the question is can and should these new entities have rights? This question has been put to the rights community most succinctly by David Gunkel (2018), and we have yet to rise to the challenge. This debate will likely be animated by old rivalries, as natural rights proponents square off against constructivists. There is reason to hope, however, that the technologies described here add complexity to the conversation, as one can imagine a natural rights proponent arguing for the rights of an entity that has resulted from whole brain emulation, should the copying of a mind ever become possible (specialists are extremely skeptical). It is also likely to attract attention from religious communities, as evidenced by recent studies on transhumanist communities in Russian Orthodoxy (Bernstein, 2019) and American Mormonism (Bialecki, 2022).

Expanded Agenda—The human rights agenda should be expanded to anticipate opportunities, rather than just respond to threats. Changes to the human—the erstwhile human rights movement’s sole unit of analysis—are only one part of this change. There are many others, including the rise of China, the onset of demonstrable and high-impact climate change, and the shift from oil to renewables (and the subsequent redistribution of capital). Taking the future seriously would involve expanding the rights agenda at every level, but especially within the intergovernmental institutions that set the tone for larger debate and agenda for subsequent resources (what Steven Hopgood has referred to as “Human Rights”).

New Goals—The human rights field’s focus on Homo sapiens needs to be radically updated. Innovation in the policy space is according rights to nature and nonhuman animals, and innovation in the technology space is raising provocative questions about the rights of superintelligence and social robots (Schulz & Raman, 2020). The rights of Homo sapiens must be protected, but creative and proactive work is needed to clarify who else counts among the deserving and whether the rights framework is the right tool.

See, for example, Choi-Fitzpatrick (2015, 2016, 2017), Brysk and Choi-Fitzpatrick (2012), Choi-Fitzpatrick and Watkins-Smith (2021), and Boyd et al. (2021).
for the job. At the most fundamental level, this effort should begin by engaging scholars doing boundary work in the area on flourishing. These include Donna Haraway (2016), who advocates for making kin in a way that advances multispecies flourishing, and Martha Nussbaum (1998), who has specified a number of possible paths for supporting multispecies flourishing.

**New Interlocutors**—The human rights agenda must make common cause with newly discovered fellow travelers. Lawyers, policymakers, advocacy groups and activist networks, heads of state, and hashtag campaigns are vital. They should be joined by artists, engineers, computer scientists, science fiction writers, and a host of other sympathetic fellow travelers in the natural sciences, practical arts, and humanities. If we want to see new things, we need new perspectives, new frames of reference, and new epistemic communities.

**New Tools**—The human rights toolkit needs to be radically updated. Policy reports and advocacy campaigns are important and effective. They should be joined by the tools brought by new interlocutors. These are likely to include established methods like worldbuilding, rapid prototyping, and random controlled trials alongside the imagining and development of experimental approaches.

**Expanded Resources**—The future needs to be an appropriately resourced component of the human rights agenda. While it is important to have think tanks and international organizations primed to respond to urgent challenges of the hour, it is also important to follow the advice of the prophets who would have us set a person on the watch, to declare what they see. Resources for established and new future-focused organizations and networks should be scaled up significantly.

**New Coordination**—Work across all these areas (agenda, goals, interlocutors, tools, and resources) must be tracked, and perhaps coordinated, at the highest level. The American professor of geology Marcia Bjornerud has suggested the establishment of a national Department of the Future. Science fiction writer Kim Stanley Robinson has upscaled that idea to suggest a Ministry for the Future, in a book by the same name. The idea, at least philosophically, traces back to Kurt Vonnegut, who suggested the establishment of a Secretary of the Future, perhaps mindful of seven-generation stewardship practiced by both the Iroquois Nation and the adherents to the Jewish faith.

A Ministry for Future Rights may be too ambitious of an objective, but an ambassador-level position at the United Nations seems within reason. Such an office would coordinate with advocacy groups, businesses, national governments, and major institutions at the national, international, and transnational levels. The number of actors and size of the ecosystem is growing, but the international—and human rights—coordination capacity does not appear to have kept pace.

What, then, is to be done? Scholarly debate, legislative efforts, social movements, and active judiciaries each do a wonderful job of tending to and occasionally refining or reforming the status quo. An expanded agenda, new goals, new interlocutors, new tools, expanded resources, and new coordination are needed, however, if a revolution is on the horizon.

**IX. Conclusion**

The future is uncertain. It may indeed be the end of the road for a certain instantiation of human rights, but this critical juncture is more of an opportunity than it is a dead end. By organizing this essay around the historian-infuriating concept-clarify-consolidate rubric, I hope to have set us up for a fresh conversation about what gets invented, clarified, and normalized as rights in the future.

The argument advanced here is that the present moment is neither here nor there. Such a space can be thought of in terms of **liminality**. The Belgian anthropologist Charles van Gennep (2013) coined this term, drawing on the concept of the **margin**—the in-between space lying where medieval European city-state sovereignties didn’t quite meet. It is this same moment Gramsci was referring to in observing that “the old is dying and the new has yet to be born.”

One thing is clear—Independently, but especially altogether—we are entering an era of significant challenges to the normative order that the current political economy is able to sustain. If anything, then, this essay is a draft map of the unknown, sketched with the hope of driving scholarly debate and inquiry.

Futurology is foolishness. It is similarly shortsighted to assume that the most significant challenges facing the human rights enterprise are to be found in the present. There is no way to predict the future. This does not mean that we should not debate it, anticipate it, and work to shape it. Marx’s maxim that “[humans] make history, but not under conditions of our own choosing” should prompt debate about how we understand the conditions that we operate in. My efforts here suggest that human rights scholars and activists would be well served to start asking questions about how big changes may alter the space within which our efforts take place, the raw material we have to work with, and the entities our efforts are intended to benefit.
Works Cited


Bryson, J. J. (2010). Robots should be slaves. Close Engagements with Artificial Companions: Key Social, Psychological, Ethical and Design Issues, 8, 63–74.


Statements and views expressed in this report are solely those of the author and do not imply endorsement by Harvard University, the Harvard Kennedy School, or the Carr Center for Human Rights Policy.

Copyright 2022, President and Fellows of Harvard College
Printed in the United States of America