Human Rights in the Big Data World

Francis Kuriakose\textsuperscript{1} and Deepa Iyer\textsuperscript{1}

\textsuperscript{1}Affiliation not available

February 19, 2019

Abstract

Ethical approach to human rights conceives and evaluates law through the underlying value concerns. This paper examines human rights after the introduction of big data using an ethical approach to rights. First, the central value concerns such as equity, equality, sustainability and security are derived from the history of digital technological revolution. Then, the properties and characteristics of big data are analyzed to understand emerging value concerns such as accountability, transparency, tracability, explainability and disprovability. Using these value points, this paper argues that big data calls for two types of evaluations regarding human rights. The first is the reassessment of existing human rights in the digital sphere predominantly through right to equality and right to work. The second is the conceptualization of new digital rights such as right to privacy and right against propensity-based discrimination. The paper concludes that as we increasingly share the world with intelligence systems, these new values expand and modify the existing human rights paradigm.
Human Rights in the Big Data World

Francis Kuriakose & Deepa Iyer

Abstract

Ethical approach to human rights conceives and evaluates law through the underlying value concerns. This paper examines human rights after the introduction of big data using an ethical approach to rights. First, the central value concerns such as equity, equality, sustainability and security are derived from the history of digital technological revolution. Then, the properties and characteristics of big data are analyzed to understand emerging value concerns such as accountability, transparency, traceability, explainability and disprovability. Using these value points, this paper argues that big data calls for two types of evaluations regarding human rights. The first is the reassessment of existing human rights in the digital sphere predominantly through right to equality and right to work. The second is the conceptualization of new digital rights such as right to privacy and right against propensity-based discrimination. The paper concludes that as we increasingly share the world with intelligence systems, these new values expand and modify the existing human rights paradigm.

Keywords: Human rights, Big data, Privacy, Automation, Artificial intelligence, Technology
I. Introduction

‘We can only see a short distance ahead, but we can see plenty there that needs to be done.’

Alan Turing (1950)

‘It is far better to foresee even without certainty than not to foresee at all.’

Henri Poincaré (2014)

‘Success in creating effective AI, could be the biggest event in the history of our civilization. Or the worst [...] we cannot know if we will be infinitely helped by AI, or ignored by it and side-lined, or conceivably destroyed by it’.

Stephen Hawking (2017)

The history of modern human rights is based on the valuation of the individual, which is closely related to the history of liberalism as an ideology (Hobbes 1652; Locke 1690; Paine 1792). The development of ‘Rights of Man’ was the product of a historical time that ascribed value to the economic, political and military use of individuals to consolidate nations (Hitchens 2009). Later, in the twentieth century, liberalism aligned with capitalist ideology to bridge the growing dissonance between redistribution (ethics) and production (economic growth) by introducing the concept of ‘human rights’. Since then, human rights have gone through three generations of rights—civil and political rights, socio-economic rights and the right to self-determination.

The coming of big data and artificial intelligence poses two specific threats to the liberalist paradigm because of the decoupling of intelligence from consciousness (Harari 2015). By doing so, human beings lose their value as individuals because many fundamental human activities will be efficiently performed by machines. Moreover, a select group of human beings will find themselves relevant and powerful in the age of automation because of the capital they hold—skillset, wealth, or knowledge—leading to the formation of a class of digital élite. These two tendencies will invariably lead to inequality. Regarding its impact on human rights, there are two concerns of interest here. The first is the extension of existing human rights and its adoption in the digital sphere. The second is the creation of human technological rights that are highly specific to the infosphere. For both these deliberations, it is important to understand the premises on which existing human rights have been conceived. We can anticipate an institutional arrangement that ranges from techno-humanism (where human rights based on liberalism is extended to the digital sphere), dataism (where value of any person or entity is determined by their contribution to data production or processing) or a new paradigm of rights.

In this context, this paper examines human rights in the big data world using an ethical approach. The ethical approach goes beyond the framework of law and understands the normative paradigm that give rise to rights. For this, the paper evaluates the central value concerns of digital technology
by using a historical review of technological pathways. Then, the properties of big data are examined to understand new and emerging value concerns. This paper anticipates the future typology of human rights from these value concerns. We propose that big data calls for a re-evaluation of existing human rights in the digital sphere as well as the introduction of new digital rights.

II. Origin and Nature of Human Rights

Human rights are claims of rights made by people on others (Luban 1980). Therefore, human rights necessitate a class of beneficiaries and obligates. Furthermore, a human right is a basic right that is necessary for the fulfillment of other rights. The idea of defending human rights is based upon certain assumptions (Lukes 1993). First, human rights are conceived as ‘restraints’ on individual liberty to enable the realization of some collective good. Second, human rights invoke a certain level of abstraction from the local and the specific to the more universal aspects of life. Third, the notion of human rights presupposes a set of permanent existential facts about the human condition. The modern notion of human rights began within the liberal tradition. The precursors of human rights were ‘inalienable rights’ in the Bill of Rights of American Independence, ‘Rights of Man’ of the French Revolution and ‘universality’ of rights adopted by the United Nations Declaration of Human Rights. The liberal tradition, that was founded on the centrality of individual will and liberty has been instrumental in the making of generations of human rights from civil and political rights (first generation), economic and social rights (second generation) and right to national self-determination (third generation). The extension of existing human rights into the infosphere and the emergence of new rights in the age of big data can be seen as the ‘fourth generation’ of rights in the human rights paradigm.

The concept of human rights takes meaning not only from the manner in which such rights are realized in practice today, but also the historical context in which those rights were conceived. This approach to human rights acknowledges the legalized aspects of rights (what can be claimed by law) as well as the ethical principles from which they draw upon (what ought to be the basis of law). Following this understanding, Sen (2009, p. 358) argues that ethical assertions behind human rights have two underlying features. The first feature is the ‘content’ of rights and the second is the ‘viability’ of rights. The content of human rights denotes a set of freedoms given with a corresponding set of obligations. The viability of human rights indicates the ethical principles behind rights that has to be examined with open and impartial reason.

---

1 Luban (1980) discusses basic security rights and subsistence rights as two types of human rights. This identification is closely related to the Hobbesian dependency on state for the realization of these rights.
2 For a discussion on the early definition of human rights, see Cranston (1964). For an evaluation of rights in the liberal tradition, see Shapiro (1986). On the insufficiency of the liberal notion of human rights, refer Farer (1985) and for a comparative perspective of various traditions of rights, see Pollis (1982).
3 For a comprehensive discussion on the liberal roots of modern human rights including generations of rights, refer Howard & Donnelly (1996).
4 For a detailed discussion on the generations of human rights, see Ishay & Goldfischer (1996).
Beyond the components of ethics, the relationship between ethical assertions and legal rights are important when we discuss the emergence of new contexts such as big data. The interconnectedness of ethics and law has been analyzed by two opposing schools of thoughts in philosophy. The first is the utilitarian tradition of Bentham that sees ‘rights as child of law’, i.e., the belief that rights come as a result of legal enactment (Bentham 1792). This tradition distinguishes between the actual legal status of ethical declarations and legislated rights. The second tradition is in the vein of Paine’s Rights of Man that advocated ethical norms as the basis of legal rights (Hitchens 2009). Other legal philosophers such as Hart (1955) have argued that rights are derived from a branch of morality and is appropriately decided for legislation. In effect, the second school of thought conceives rights as ‘parents of law’.

The ethical roots of human rights indicate how ethics is related to law in practice. The legislative route through which human rights become law is important for its design, monitoring and implementation (Sengupta 2000; 2004). Such an understanding is also significant for amending the legal institutions through deliberative processes such as public scrutiny, discourse, education and media responsibilities. If the principles of ethical considerations behind human rights are well laid out, the evaluative framework (content and viability) of human rights can be reasonably assessed. Sen (2009) discusses the variables of the evaluative framework that includes opportunity and process aspects of rights, perfect and imperfect obligations as well as interests and freedoms of stakeholders.

Furthermore, understanding the ethical foundations of human rights is also important (i) when the existing rights are reviewed for new circumstances, and (ii) when new rights are deliberated upon. The main focus of this paper, which is to understand the implications of big data on human rights, calls for such a deliberation. Before we analyze the impact of big data on human rights, it is important to understand how technology evolves and impacts our lives in various ways. In the next sections of the essay, we discuss the nature of technological innovation pathways including the waves of digital revolution, the characteristics of big data and the cultural paradigm big data has brought in.

III. Development of Innovation Pathways

The manner in which any type of technological innovation diffuses in society is determined by innovation pathways. The study of history of technology is useful in understanding how it is likely to unfold in the future and impact the society. This is because the pattern of technological pathways repeats over short and long timescales from which it is possible to gauge how specific type of technological products and designs might be received in the world. This section discusses how

---

5 Dasgupta (1982) discusses the relation between information and rights using utilitarian principles to derive the principles of distributive justice. For a recent discussion on the moral basis of law, see Raz (2010).
6 Sen (2009) conceives rights as manifestations of freedom conditioned with obligations. He distinguishes the actual opportunities (ends) accorded by the rights and the process (means) through which they are realized.
technology grows from the stage of conception to widespread use in the society and how digital technology evolved through the last five decades.

i. **Technology Diffusion Curve**

Innovation pathways follow an S-shape curve consisting of four distinct stages- pre-development, initial take-off, acceleration and stabilization phases (Rotmans, Kemp & van Asselt 2001). In the pre-development phase, a new technology is introduced in specific subcultures of the society who have a specialized need for them. The design and use of early versions of a new technology are entirely determined by this subculture that is subterranean. In the second and third phases, the adoption of technology accelerates either because two or more technologies combine to become viable or a critical mass of people adopt the technology, leading to a vertical take-off. Finally, technological adoption matures in a stabilizing plateau when it is adopted *en masse* and levels off. Figure 1 depicts the typical innovation pathway curve.

**Figure 1**

*Innovation Pathway Curve*

Other than historical analysis, the second insightful avenue to understand the technology of the future is the ‘edge cases’ of the present (Standage, 2017, p. 13). New technologies and behaviour associated with them have a pattern of remaining popular with specific sub-populations for a

---

7 For a discussion on how technological subcultures lead the evolution of technological diffusion, refer Standage (2017).
period of time before going mainstream. Therefore, today’s popular technologies have specific design and functional default settings that they imbibed from the time they were edge cases. For example, the quantified self-movement, that tracks the health and fitness levels of people through tracking devices, had a cult status among health aficionados before it became a more widespread technology.

Finally, the futuristic vision provided by scientific fiction culture also indicate how contemporary ideas and present-day concerns of technology entrepreneurs shapes the future. Elon Musk, the founder of the automobile manufacturer Tesla and an American venture capitalist, calls the impact of science fiction ‘branching probability streams’ from which technology entrepreneurs get inspiration as well as signals of caution (Standage, 2017, p. 15). Perspectives from fictionalized representations are valuable in evaluating problems such as overdependence on machines and environmental destruction.

ii. Waves of Digital Revolution

Beyond innovation pathways, digital technological progress is specifically determined by empirical observations of regularity that have received the status of ‘laws’. For example, Moore’s Law describes how computing power evolves over time. Moore’s Law states that the number of components that could be incorporated into an integrated circuit double every year (Moore 1965). Furthermore, Dennard’s scaling noted that shrinking of a chip’s size made it cheaper, faster and less power hungry (Dennard et al. 1974). Since the commercial availability of the world’s first microprocessor called 4004 in 1971, there have been 22 rounds of Moore’s Law in action (Cross 2017). The reliability and repeatability of this pattern has been instrumental in short product cycles and shrinking size of components in the computing industry for the last five decades.

The understanding of the history of digital innovation pathway and technological generations through ‘waves’ is partly due to the metronomic regularity of the Moore’s Law (Baldwin 2016; Cross 2017).

---

8 The quantified-self movement (QSM) is an example of how data has influenced epistemology, that is discussed elsewhere in the paper. QSM was conceived by health aficionados who measure and track every aspect of their body to learn new things about themselves.

9 Other technological ‘laws’ derived out of Moore’s Law are Gilder’s Law and Metcalfe law. Gilder’s law states that bandwidth grows at least thrice as fast as computing power. Metcalfe law states that the effect of a telecommunications network is proportional to the square of the number of connected users of the system. For more, see Baldwin (2016).

10 In 1971, Intel built chip 4004 from 2300 transistors, each 10,000 nanometers in length. A transistor is an electronic switch that provides the physical representation of the fundamental particles of information through its flipping on and off.

11 Scholars are of the opinion that we have reached the physical and financial limits of the Moore’s Law as it is increasingly becoming difficult to shrink the size of the chip. The computing industry has come up with several alternatives to keep Moore’s Law relevant with concepts such as 3D chip and quantum computers which require specialized care and data feeding which are currently expensive. Alternatives to Moore’s Law comes through options such as ‘cloud computing’ that enlarges the hardware that does computationally heavy task. Another concept is ‘internet-of-things’ that embeds chips in regular objects to introduce data collection and analytics by every-day objects. The bottom line of moving away from Moore’s Law is moving away from augmenting computational power per se. For a detailed discussion, see Lanier (2010), Mayer-Schönberger & Cukier (2017), Cross (2017).
The seven waves of the digital revolution including the core technology they introduced are discussed in table 1.

<table>
<thead>
<tr>
<th>Digital Wave</th>
<th>Core Technology</th>
<th>Impact</th>
</tr>
</thead>
<tbody>
<tr>
<td>One</td>
<td>Mainframe and minicomputers</td>
<td>Mainstreaming of hardware manufacturers</td>
</tr>
<tr>
<td></td>
<td></td>
<td>such as Intel</td>
</tr>
<tr>
<td>Two</td>
<td>Desktop software and operating systems</td>
<td>Introduction of software industry such as</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Microsoft</td>
</tr>
<tr>
<td>Three</td>
<td>Internet through Web 1.0</td>
<td>Internet-based commercial firms such as</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Amazon and Google</td>
</tr>
<tr>
<td>Four</td>
<td>Cloud and mobile computing through Web 2.0</td>
<td>i-Phone and Android</td>
</tr>
<tr>
<td>Five</td>
<td>Big data, analytics, visualization</td>
<td>Open sources codes</td>
</tr>
<tr>
<td>Six</td>
<td>Internet-of-things, smart machines</td>
<td>Autonomous cars, drones</td>
</tr>
<tr>
<td>Seven</td>
<td>Artificial intelligence</td>
<td>Machine learning and deep learning</td>
</tr>
</tbody>
</table>


There are three distinguishing features that have stayed consistent throughout the seven-wave growth of the digital revolution. The first is that each wave has grown exponentially benefitting from the multiplier effect of the core technology of the previous wave. For instance, operating systems and desktop software that were the core technology of the second wave began as a method of solving the problem of large hardware from the first wave. The second feature of a digital wave is that the pace of growth has increased exponentially as each wave matured. For example, while Microsoft took 15 years to cross the $1 billion revenue mark, Google reached the same milestone in five years (Winblad, 2017, p. 68). The third feature of the digital wave is that each leap in core technology has brought with it a characteristic business model that has commercialized the technology. A typical example is the coming of venture capitalists to fund the expansion of the software industry in the second wave.

The historical view of technology helps us understand some of the fundamental and underlying principles on which technological substrate builds on the society-at-large. We now focus on the specific characteristics of big data to understand the scope and extent of its penetration.

IV. Big Data: Properties and Limitations

Data refers to a description of something that can be measured, recorded, analyzed and re-organized. To make information into data is therefore to quantify it so that it could be tabulated...
and examined\(^{12}\) (Mayer-Schönberger & Cukier, 2017, p. 78). The process of datafication (transforming information into data) enables us to replicate human activities, predict trends, plan events, correct errors in estimates and audit decisions.

Big data refers to the society’s ability to garner information in novel ways to produce goods, services or insights of significant value (Mayer-Schönberger & Cukier 2017). Big data has transformed our idea of what can be converted into data. For example, over the waves of digital revolution, words, location, social relations and the idea of oneself have been transformed into data points through search engines, maps, social networking and quantified-self movement. Furthermore, ‘metadata’ (information about information) has also spawned new portals of data collection and analysis. For example, a single post in a social networking website can generate metadata about user profile, location, network and psychological variables such as moods and attitudes of users, that have regularity and pattern.

Big data has become a raw material and an input in business as it is able to generate value (and revenue) by applying mathematical computation to information and infer probabilities. The volume of information that is available today is very large and growing at a faster rate. For example, Google processes more than 24 petabytes of data per day that is thousands of times more than the total amount of printed materials available in the US Library of Congress (Mayer-Schönberger & Cukier, 2017, p. 8). This change of scale in the quantity of data is leading to a qualitative change in the way we understand the world. For example, big data calls for an ability to deal with very large sets of information, be comfortable with its messiness\(^{13}\) and settle for correlations rather than causations.

The reason why big data has gained traction is because of its unique properties that generate economic value. The first interesting feature about big data is that it has made all data valuable. This means that raw and processed data, structured and unstructured data, passively collected mundane data and metadata are all potential inputs for financial returns. This is in contrast with the time when data was valuable as an auxiliary input in business practices and was limited to specific categories such as intellectual property and information.

Second, the same data can be harnessed for multiple uses. This means that data can transform itself from a form that generates primary value to potential future uses that might generate secondary value. An illustration is the case of ReCaptcha\(^{14}\), a concept by which a set of two-word digital text is presented to distinguish human users from spam bots. The first of the words is intended to affirm that the user is human and the second word is digital text that needs to be deciphered by human

---

\(^{12}\) The idea of information of data (points of measurement) should be contrasted with that of information as content (points of meaning).

\(^{13}\) The messiness of big data refers to its inexactitude which is a function of imprecise tools at such large scale. For more, refer Mayer-Schönberger & Cukier (2017).

\(^{14}\) ReCaptcha was the firm that was founded by Luis von Ahn, then graduate student of Carnegie Mellon University, who originally invented the concept of Captcha to distinguish between human users and spam bots by asking users to decipher digital texts.
users for the first time. By presenting the same text to at least five human users, the same data fulfills its primary use of distinguishing humans and secondary use of deciphering new digital text.

Third, big data has use value and reuse value. In this sense, data is what we call non-rivalrous good in economics, i.e., one individual’s consumption of data does not affect its use by another. Furthermore, data does not suffer diminishing quality on recurring usage, the way material goods and other inputs of production do. For example, internet firms like Amazon repeatedly uses data based on previous transactions of one customer to recommend products and services to several others.

Fourth, big data can be harnessed inexpensively and passively\textsuperscript{15} as long as there is an appropriate internet infrastructure. Furthermore, the cost of data storage has fallen over the last five decades even as storage density\textsuperscript{16} has risen million-fold, resulting in low cost for data storage.

These features make big data an emerging input in production. Big data is transforming how firms view their organization and supply chain, and understand and utilize the type of data they have access to. Due to the possibility of multiple uses, data is also being seen for its potential ‘option’ value through reuse, merging of data sets and recombinant datasets\textsuperscript{17}. Big data is rising as an important input of production as well as a firm asset. An example is the evaluation of Facebook’s worth through its book value of $6.3 billion and market estimate as $104 billion during its initial public offering in 2012, as most of Facebook’s asset is in ‘data form’ (Mayer-Schönberger & Cukier 2017). Furthermore, the big data supply chain has also created three types of big-data firms—data producers, data analyzers and data-mindset companies as well as a host of data intermediaries. Data producers are those firms which generate data or have access to it, data analyzers analyze data and extract value out of it and data mind-set firms spot the potential value of data and commercialize it.

The features of data have also led to certain design principles by which firms extract the maximum value from the data that it collects. First is idea of ‘extensibility’, or the possibility of using data in different ways, is built into the design of data collection from the outset. A common example is the use of CCTV cameras in shops, whose primary use is security, but has secondary uses such as understanding the effectiveness of store layout and marketing strategies. Second is the idea of ‘data exhaust’ or data that is produced as a by-product when users navigate the internet (Mayer-Schönberger & Cukier 2017). This hidden form of data is collected and recursively used as an input in a feedback loop to learn from the data. Examples of using data exhaust range from personalizing the internet use experience, improving a service to creating an entirely new digital

\textsuperscript{15} By passive, we mean, data can be harnessed without much effort, intention or awareness on the part of users or firms generating it.

\textsuperscript{16} Data storage density (areal density) is a measure of the quantity of information bits that can be stored on a given length of track, area of surface, or in a given volume of a computer storage medium.

\textsuperscript{17} Recombining datasets is a method of commingling data that were previously discrete to form interesting new combinations.
product. An example is Google spellcheck that not only helped to correct spellings but has introduced autocomplete based on past data patterns.

The new role of big data in human transactions and its evaluation as a new valuable commodity and input in production has directly and indirectly impacted the content and essence of our cultural paradigm. This is discussed in the next section.

V. Big Data’s Cultural Paradigm

Big data technology modifies our functioning in this world through a paradigmatic shift introduced by three types of changes in our cultural environment. First, big data changes the way we communicate socially. In this sense, big data acts as a medium of communication. Postman (1985, p.10) describes what a medium of communication offers a culture by influencing its major intellectual and social preoccupations:

> Each medium, like language itself, makes possible a unique mode of discourse by providing a new orientation for thought, for expression, for sensibility. [...] The forms of our media, including the symbols through which they permit conversation, [...] are rather like metaphors, working by unobtrusive but powerful implication to enforce their special definitions of reality. Whether we are experiencing the world through the lens of speech, or the printed word or the television camera, our media-metaphors classify the world for us, sequence it, frame it, enlarge it, reduce it, colour it, argue a case for what it is like.

In other words, big data by acting as a medium of communication, changes the way we think and therefore presents us with the content of our culture. It gives us a world view, a new sense of audience and posterity. Lanier (2010, p. 4) argues that technology is in this sense ‘extensions of ourselves’ that engages in social engineering. For example, digital technologies demand a specific type of interaction from people. Here, a vast number of anonymous individuals conceive the content leading to our unconscious acceptance that a faceless crowd has a legitimate point of view, obscuring aspects of ‘pack mentality’ that suppress individually reasoned arguments. It is important to realize that nothing about digital technology - the structure and ambience, design elements such as user interface and the way data is managed and valued - is inevitable. All the aspects that we take as ‘default settings’ were chosen out of a plethora of choices available to us at the point of conception and development of internet services. This perspective shows us that the openness and radical freedom of the infosphere can be associated with self-restrained use and a personal sense of accountability before thinking about institutional regulations.

The second way in which data technology changes our functioning in the world is by presenting us a novel epistemology, what Postman (1985, p.12) calls ‘a new conception of knowledge and a new sense of intelligence’. By performing this role, technology influences the origins and nature of knowledge and definitions and sources of truth. Postman uses Frye’s concept of ‘resonance’ to explain how the mechanics of creating new epistemology works (Frye 1981). Resonance is the

---

18 While technology creates new ways of social behaviour, we do not espouse the opinion that technology has a will of its own that humans cannot resist or regulate. In fact, we distinguish the idea of ‘technium’ (discussed elsewhere in the paper) with that of political philosophy of self and institutional regulation. For a discussion on the different perspectives on this issue, see Kelly (2010); Lanier (2010).
process by which a metaphor becomes a generative force, i.e., a particular statement, object or idea in a particular context assumes universal significance. In other words, a new medium changes the structure of public discourse by favouring certain uses of the intellect, definitions of intelligence and by demanding certain kind of content (Postman, 1985, p. 27). The changes made in the intellectual environment is gradual and additive till a critical mass of people subscribe to the new view after which it becomes more widespread.

The third manner in which data technology impacts society is the way it transfers labour from humans to machines, adding material value and ushering economic growth. It is important to note not only what machines accomplish but how they do so. This is relevant in understanding the digital revolution that created computers to smart machines through progressively transferring specific types of intelligence and abilities onto machines. Moravec\(^{19}\) (1988) has distinguished machine intelligence from human’s using the axiom that machines have become adept in replicating computational power and analytical ability (aspects of human thought) rather than replicating sensory perceptions and mobility (aspects of human movement). The result has been a steady rise in living standards, material comforts and improvement in human rights (Kasparov 2017).

These three aspects of change - new forms of social interactions, new understanding of knowledge and progressive transfer of labour - offer us the foundations on anticipating the type of human rights questions that would be of interest in the big data world. The next section discusses a framework of understanding modification to existing human rights and the emergence of new rights due to the coming of the big data age.

**VI. Big Data and Human Rights: A Framework**

Big data poses interesting challenges to traditional regulatory institutions that have been built on the assumption that information is limited. Even the information that institutions use to make estimates are from samples that lack granularity and real-time dimensions. But today, there is more data than ever and more participation from interconnected users even though the information given by the data is not better or more representative. This poses specific opportunities and challenges to human rights.

McPherson (2015) identifies the opportunities that internet and communication technology (ICT) presents to the advancement of human rights. Whilst ICT offers faster mode of communication, early warning systems, opportunities for communications-based conflict resolution and quick pathways for mobilizing public discourse and advocacy, the same internet space also presents new and increased security threat to human right defenders and brings new types of evidence from unknown and unnamed sources. Furthermore, the internet divide between those who can access and effectively use the medium and those who are locked out of it results in new types of

---

\(^{19}\) This inference is described as Moravec’s paradox, i.e., machines (computers) are adept in performing activities that humans are weak at and vice versa. This was deduced by robotist Hans Moravec on comparing computer’s performance with adult intelligence.
inequalities. The existing legal frameworks and principles (refer Appendix I & II) are constantly in need of revision and consultation.

In this context, Guberek and Silva (2014) note that there have been efforts to extend existing human rights in the digital space and acknowledge the existence of new rights by various types of international organizations\textsuperscript{20}. The main tenets evolving from these consultations are (i) the necessary and proportionate principles for adjudicating crime and punishment in the digital realm, (ii) monitoring and reporting of human rights violation on the basis of evidence, (iii) cross networking between human rights defenders in the physical world and those in the digital world\textsuperscript{21}, and (iv) improved data management to confront new concerns regarding data security\textsuperscript{22}.

From the historical analysis of technological pathways and specifically data technology, we identify the value concerns that come from the characteristics of data technology. This framework is then used to understand the types of human rights concerns big data evoke.

i. Value Concerns for Existing Technological Rights

Technology could be gained as well as lost in the course of time. This fact alerts us to the possibility of technological dead-ends as much as exponential growth. Technological change could therefore evolve either in an inter-generationally stable or unstable way. Whenever the scale and rate of production changed with technology as in the industrial revolution of the eighteenth century, the world inhabited by the next generation was completely remodeled from that of the previous generation. The value concern here is that of sustainability.

Second, technological pathways are sometimes interpreted as having a will of their own. Kelly (2010) calls the sum of all things technological ‘technium’. He argues that ‘technium’ defines technological progress as increasing connectivity between people and complexity in design under the assumption that people benefit materially and spiritually from such an arrangement. However, scholars like Morton (2017) caution against such an approach of treating technology as power unto itself. He argues that if the will of ‘technium’ diverges with that of humankind, it would be the humans who have to realign their interests with technology. This is because infosphere shapes our physical and conceptual environment and constraints our abilities to adapt to it (Floridi 2017). Several scholars have argued similarly that ‘encapsulation’ in an artificial intelligence (AI) friendly environment would make machines more at home than humans (Lanier 2010; Floridi 2017). In this context, it is also important to remember that we share the infosphere with other digital technologies who are computational machines with the need to be fed with copious amount of data. As Floridi remarks, ‘no conscious, intelligent, intentional entity is magically going to

\textsuperscript{20} The main sources of evolving human rights norms in the digital sphere are the UN Special Rapporteur on the Right to Freedom of Opinion and Expression (2011), a group of 18 national governments called Freedom Online Coalition, UN Guiding Principles on Business and Human Rights (2011) and International Principles on the Application of Human Rights to Communication Surveillance. For more, see Appendix I & II.

\textsuperscript{21} For a detailed discussion on extending human rights in the physical world to the digital sphere, refer Zuckerman (2013).

\textsuperscript{22} Evaluation projects with narrow objectives and short life cycle need to be reassessed in this context.
emerge from a Turing machine’ (Floridi, 2017, p. 160). The value concern here is equality of human beings in a world that we share with digital technologies.

Third, the history of technological innovation also shows the central role of capitalism in its adoption and expansion. The reinvestment of capital in market-driven growth creates large-scale demand for and deployment of technology that is unlike any other mode of organization23. For example, in industrial revolution, a culture of innovation through mechanical improvement was encouraged through systematic re-investment of capital that increased the output exponentially and lowered the cost of production (Malm 2016). The central value here is the concern for equity that comes out through how to redistribute benefits of the investment more widely with the world.

Finally, historical view of technological innovation alerts us to externalities or unintended consequences (Morton 2017). As innovation progresses, there are unknown variables and combinations that emerge from the interconnectedness of the systems that can be destructive to the environment, social relations at the workplace, or psychological well-being at the individual level. With augmented memory, expanded sensory perceptions, increasing connectedness and possibilities of creating other realities, the digital revolution has unleashed in new categories of negative externalities. The central concern here is the security aspect of identity and safeguard against manipulation and discrimination. The roles of state and non-state actors are important in such a discussion.

Figure 2

Technology and Value Concerns

Source: Authors’ Compilation

---

23 The analysis of capitalism and its role in technological innovation does not preclude the central role of governments in large-scale investments to promote and diffuse technology in the twentieth century.
These four historical features of technological evolution have a direct bearing on human rights because they show the central value concern at every level (Figure 2). Understanding these conflicts would help us adopt the most optimal technology for our society and also safeguard our rights.

ii. Old and New Human Rights

In this context, we propose a framework to examine sets of human rights concerns that big data evokes (Figure 3). Human rights discourse in the big data world deals with (i) existing rights that need to be extended into the digital sphere as well as (ii) new digital rights. Within existing rights, right to equality in the infosphere and right to work and welfare need to be reworked to fit in with new values and challenges. Within new digital rights, right to privacy and right against propensity-based discrimination emerge as new concerns that need to be addressed.

**Right to Equality in Infosphere:** The first set of questions in extending existing rights is regarding equality of data flow and access. Net neutrality is the principle by which all data flows are given equal priority. Lanier (2010) argues that formulating such principles is important in resisting privileged internet services and selling online labour of many for the surveillance and advertising benefits of the few24. In effect, aspects of design that determine data flow and access deal with digital rights of all users and the type of organizational model that governs the internet. The issues here are that of access and control between the stakeholders - users, volunteers, for profit business firms. Studies have assessed the existing types of internet governance based on the regulation of access as open commons, access denied, access controlled and access contested forms (Deibert et al. 2011). The issue of disparity in institutional strength and technical capacity of various stakeholders such as data producers, collectors and analyzers, leads to a new set of adversaries and interlocuters in regulation (Guberek & Silva 2014). The human rights concerns of equality and equal access is directly relevant in this context.

---

24 It is important to understand that the concept of net neutrality principle is inherently opposed to the ‘open ideal’ of the internet. For discussions on net neutrality, refer Economides (2008), Kramer et al. (2013), Yoo & Srinivasan (2017).
Right to Work: The second set of questions that come in modifying existing human rights is regarding the comparatively limited transformative potential of the digital revolution in terms of work and wages. As Avent (2017) argues, digital revolution has not improved the output per person, life expectancy or wage levels as much as industrial revolution and the consequent improvement in welfare did. This is because technology takes time to mature and transform productivity, often going through long periods of dormancy before accelerating. During this period, the improvements from technology is only marginal. Moreover, the value produced in service and information industry is not comparable to the value added in output through the previous industrial revolutions. In data technology, the value added is due to the rising quality of personalization and is hard to measure.

These effects have direct impact on the way work-wage relationship is conceived (Baldwin 2016; Avent 2017). For example, wage linked to performance that is the hallmark of twentieth century labour will be irrelevant as jobs are increasingly automated. Second, increasing automation and introduction of artificial intelligence result in a class of subservient service class labourers. Questions on how to organize our social system to distribute the benefits of automation more widely will increasingly come up. This will involve consideration of higher wage subsidies, unconditional universal basic income and other programs for which a higher and wider tax-base is
necessary. Furthermore, society needs to adapt to a situation where work becomes optional as a source of income.

Other than reassessing existing rights to make them relevant for the digital sphere, new digital rights also emerge and bring with them, new values. Right to privacy and right against propensity-based discrimination are two such rights.

**Right to Privacy:** The idea of personhood is challenged by big data that collects personal information. One specific concern within personhood is that of privacy. The question of privacy encompasses (i) the right to erasure such as the right to remain anonymous or be forgotten and (ii) the right to be excluded from surveillance, targeting and censorship. As Lanier (2010, p. 200) argues, strategic forgetting is part of personhood, i.e., to be a person one has to be real and inventive at the same time. This involves a nuanced balance of self-expression and appropriation in the digital sphere.

The traditional regulatory response to privacy has been through notice and consent, anonymization and opt-out preferences (Mayer-Schönberger & Cukier 2017). In the case of big-data, all these responses fail to be effective because data is being collected in large amounts, used longer and reused repeatedly. In this scenario, privacy regulation will shift from individual user consent to data user accountability (Mayer-Schönberger & Cukier 2017). Regulatory authorities will ask data user to provide (i) ways in which personal data is intended to be utilized, (ii) risks identified and (iii) safeguards put in place.

In effect, this would mean establishing principles of assessing risks, adaptability and mitigation. Legal liability and penalty measures from fines to criminal prosecution will be framed based on the new assumptions. The length and uses to which personal data is used will depend on the latent risks identified in data and social values that shape how data can be used. Holding data users accountable also rests on fair principles since they have more information about the intent of data use and stand to profit from data value. This question has direct bearing on data collection, storage and type and time-period of use.

**Right against Propensity-based Discrimination:** The question of propensity includes the right against discrimination due to the use of profiling, guilt by association and predictive analysis. The implications of using big data analysis by regulatory authorities is already apparent in the way policing is done and courts work in many jurisdictions. For example, in the US, several state police groups such as California, Washington, South Carolina, Alabama, Arizona, Tennessee, and Illinois use ‘predictive policing’ as a practice by which data analysis based on algorithms predict individuals, sub-populations and neighbourhood where criminal activities are likely to happen.

---

25 Refer Appendix III for a list of privacy laws in various national jurisdictions.

26 For a discussion of using data for predictive policing, refer Greengard (2012), Walker (2014), Mayer-Schönberger & Cukier (2017). It is important to note that the idea of predictive policing existed even before the widespread use of data. ‘Broken window’ hypothesis by which the external manifestation of a neighbourhood was used to predict its criminal tendencies was widely used by New York Police Department.
Similarly, Crime Detection Utilising Statistical History (CRUSH) is a program used by police department in Memphis, Tennessee to scout patrolling in specific areas of interest in select time periods during the day. Furthermore, Future Attribute Screening Technology (FAST) uses data analysis to detect individual physiological behaviour and pattern such as body language and vital signs to predict their intent to do harm.

From preventing ‘crime’ with probabilistic tendencies given by big data, penalizing and punishing perpetrators is not far away. The idea of detecting criminal attributes before an actual action takes place violates several fundamental jurisprudential principles such as the assumption that ‘innocent until proven guilty’. Beyond criminal justice, the use of big data to predict human behaviour can have similar consequences in all instances of human judgment aided by predictive analysis. In effect, it takes away the notion of choice and free will, by which individuals are held responsible for their action.

The traditional regulatory response to propensity-based surveillance is to have independent monitoring authority that assesses the performance of security agencies. But in the cases where both the state and commercial establishments use these tools for security and financial reasons, these measures are ineffective. Big data impacts the principles underlying the current law and order mechanisms. First, evidence-making is likely to rely on a combination of big data analysis and traditional authority to establish criminal liability. Furthermore, the data-set bringing out correlations must be (i) open and accessible to all the parties concerned, (ii) certification of the quality of algorithms used to run the data analysis as well as the quality of data used, and (iii) making transparent the limitations of the findings. This requires the expertise of data scientists to work closely with any type of data analysis that has incriminatory quality.

Thus, from the analysis of new digital rights such as right to privacy and propensity, the emerging value concerns are (i) disprovability or the ability to clearly state the limitations of final inference, (ii) explainability or the ability to illustrate how the data worked to reach the correlation, (iii) tracability or the ability to identify the source and pathway by which data was collected and retrieved, (iv) accountability or the level of confidence with which the inference can be attributed to the data and (v) transparency or the revealing of the process by which the inference was reached. The new value concerns are given in Figure 4. These principles call for a new kind of expertise and set of institutions to work with the challenges set out by big data.

---

27 Mayer-Schönberger & Cukier (2017) discusses types of data scientists of the future such as algorithmists.
VII. The Way Forward

This paper examined human rights after the introduction of big data using an ethical approach to rights. Ethical approach to rights uses normative principles revealed through values as the basis of conceptualizing human rights. Such an approach goes beyond the limits of law, often conceptualizing new grounds for extending existing laws or designing new laws. Using a historical review of digital technology, value concerns that were predominant from the evolution of big data were evaluated first. The central values were equity, equality, sustainability and security. Following this analysis, human rights concerns emerging from the impact of big data were analyzed. There were two sets of human rights—existing rights that needed modification in the digital sphere and new rights. Right to equality and work emerged as old rights that needed to be reevaluated in the light of big data concerns. At the same time, new digital rights such as right to privacy and right against propensity-based discrimination were also coming up as emerging concerns. Whilst the old values of equality, equity, sustainability and security were reflected in the new context, new values such as accountability, explainability, tracability, transparency and disprovability emerged. Human rights in the big data world is more layered and complex. An ethical approach to human rights calls for both legal constraints and normative paradigms to successfully mitigate the new challenges.

Where does the advent of big data lead human rights paradigm to? One possibility is the evolution of ‘techno-humanism’, i.e., the extension of liberal version of human rights to the digital sphere.
Techno-humanism involves extending and adapting human rights in the physical world to the digital space. It also includes conceiving new digital rights according to the principles of human rights to safeguard against violation of liberty, equality or free will. Techno-humanism works through legal statutes as well as instruments of international jurisprudence such as declarations, conventions and protocols.

As Harari (2015) has argued, are we anticipating a form of dataism as the premise of the next generation of human rights? Dataists are skeptical of human knowledge and the process by which human beings convert data to information, knowledge and wisdom. Their contention is that the overwhelming volume of data present is beyond the analytical capacity of human beings. Therefore, they prefer centralized processing methods through big data or computer algorithms. The main objectives of dataism is to increase the flow of data and enlarge the number of data points by connecting more people to the data network. In this sense, they envision a type of ‘data encapsulation’ of our digital environment and the belief that information wants to be free.

The third possibility is some form of a middle ground between techno-humanism and dataism. There are many types of permutations in this middle ground that can be feasible according to differing contexts and capacity of regulatory institutions. Understanding the value concerns and normative standards once again prove useful if we are to successfully navigate through the big data world with our rights intact.
## APPENDIX I

### UN Statutes on Human Rights/ Digital Human Rights

<table>
<thead>
<tr>
<th>Statute/Document</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Art. 7, Universal Declaration of Human Rights (UDHR) 1948</td>
<td>Equal protection against any discrimination</td>
</tr>
<tr>
<td>Art. 12, UDHR</td>
<td>Protection from arbitrary interference with privacy, family, home or correspondence</td>
</tr>
<tr>
<td>Art. 19, UDHR</td>
<td>To hold opinions without interference; to seek, receive and impart information and ideas through any media and regardless of frontiers</td>
</tr>
<tr>
<td>Art. 17, International Covenant on Civil and Political Rights (ICCPR) 1966</td>
<td>Freedom of opinion and expression with possible restrictions on the basis of rights or reputations of others</td>
</tr>
<tr>
<td>Art. 19, ICCPR</td>
<td>Freedom of expression and opinion with possible restrictions on the protection of national security or of public order or of public health or morals</td>
</tr>
<tr>
<td>UN Special Rapporteurs (UNSR) 1993</td>
<td>The right to freedom of opinion and expression in all its manifestations</td>
</tr>
<tr>
<td>The UN Guiding Principles on Business and Human Rights 2011</td>
<td>Code of conduct regarding business practices that use ICTs</td>
</tr>
<tr>
<td>UN Special Rapporteurs report to the General Assembly 2011; Human Rights Council 2013</td>
<td>Implications of States’ surveillance of communications on the exercise of the human rights to privacy and to freedom of opinion and expression</td>
</tr>
<tr>
<td>UNHRC Resolution 2012</td>
<td>Promotion, protection and enjoyment of human rights on the Internet.</td>
</tr>
<tr>
<td>UN Special Rapporteur on the Right to Privacy 2015</td>
<td>Make recommendations to ensure promotion and protection of privacy after deliberating with national and international stakeholders.</td>
</tr>
<tr>
<td>UN Resolution on Right to Privacy in the Digital Age 2016</td>
<td>To review procedures, practices and legislation regarding the surveillance of communications, interception and the collection of personal data, including mass surveillance, interception and Collection</td>
</tr>
<tr>
<td>UN Group of Governmental Experts (GGE) on Cyber Security 2017</td>
<td>Use of international law in the use of ICTs including conflict resolution</td>
</tr>
</tbody>
</table>

Source: European Union (2015)
## APPENDIX II

### Laws and Standards (Digital Rights) of Other International Bodies

<table>
<thead>
<tr>
<th>Statutes</th>
<th>Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>Art. 8 &amp; 10, European Council of Human Rights</td>
<td>Right to privacy and freedom of expression</td>
</tr>
<tr>
<td>At. 15, Convention on Cybercrime 2001</td>
<td>Fighting cybercrime; subject state powers and procedures ‘to conditions and safeguards provided for under its domestic law, which shall provide for the adequate protection of human rights and liberties’</td>
</tr>
<tr>
<td>The Organization for Security and Co-operation in Europe (OSCE) 2013</td>
<td>To enhance interstate co-operation, transparency, predictability, and stability regarding data management</td>
</tr>
<tr>
<td>Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data, amended in 2014</td>
<td>Codified data protection standards</td>
</tr>
<tr>
<td>The Internet Society (ISOC) (non-profit/non-state)</td>
<td>Set up Internet Engineering taskforce and Request for Comment protocol that made encryption a commitment for security; Developed standards and protocols form the basis of the human rights enabling infrastructure of the Internet</td>
</tr>
<tr>
<td>The Internet Corporation for Assigned Names and Numbers (ICANN) (non-profit/non-state)</td>
<td>Inputs through Government Advisory Council</td>
</tr>
<tr>
<td>The Organisation for Economic Co-operation and Development (OECD)</td>
<td>Guidelines for Multinational Enterprises on managing and regulating data in accordance with human rights</td>
</tr>
<tr>
<td>Global Network Initiative (non-profit/non-state)</td>
<td>To protect and advance freedom of expression and privacy in the ICT sector</td>
</tr>
</tbody>
</table>

Source: European Union (2015)
## APPENDIX III

**Right to Privacy in Select National Jurisdictions**

<table>
<thead>
<tr>
<th>Statue/Case law</th>
<th>Right/ Interpretation</th>
</tr>
</thead>
<tbody>
<tr>
<td>First, Fourth, Fourteenth Amendment, Bill of Rights, US</td>
<td>Right to privacy including right against unwarranted search, seizure, due process; right to protect privacy within family, marriage, motherhood and procreation</td>
</tr>
<tr>
<td>Olmstead v. United States 1928</td>
<td>Right to privacy as right to be left alone</td>
</tr>
<tr>
<td>Griswold v. Connecticut 1965</td>
<td>Right to marital privacy</td>
</tr>
<tr>
<td>Katz v. United States 1967</td>
<td>Concept of reasonable expectation of privacy</td>
</tr>
<tr>
<td>Data Protection Act 1998, UK</td>
<td>Right to privacy includes extended data protection of personal data including receiving, processing, retaining and using data in UK or anywhere UK citizen resides</td>
</tr>
<tr>
<td>Canadian Charter of Rights and Freedoms</td>
<td>Right to secure against unreasonable search and seizure</td>
</tr>
<tr>
<td>Section 21, Bill of Rights, New Zealand</td>
<td>Right to secure against unreasonable search and seizure</td>
</tr>
<tr>
<td>Data Protection Act 2012, Singapore</td>
<td>Defines personal data even if an individual cannot be identified from it</td>
</tr>
<tr>
<td>Case Law: Google Spain v. Agencia Española de Protección de Datos, European Court of Justice, 2014</td>
<td>Protection of privacy of individuals through personal data as a natural right; right to erasure</td>
</tr>
<tr>
<td>Art. 17, Regulation of European Parliament and of Council 2012; Art.17 General Data Protection Regulation 2014</td>
<td>Right of erasure and the right to be forgotten in cases where data is no longer required for the purpose for which it was collected and when the subject has withdrawn consent;</td>
</tr>
<tr>
<td>Justice Puttuswamy (Retd.) v. Union of India &amp; Others, India 2017</td>
<td>Right to privacy as a fundamental right under right to life</td>
</tr>
</tbody>
</table>

Source: Nappinai (2017)
References


