Citation: Winter, J. S. (2015). “Algorithmic discrimination: Big data analytics and the future of the Internet.” In J. S. Winter & R. Ono (Eds.)., *The future Internet: Alternative visions* (pp.125–140). Cham: Springer.

Chapter 8

Algorithmic discrimination: Big data analytics and the future of the Internet

Jenifer Sunrise Winter
University of Hawaiʻi at Mānoa
2550 Campus Road, Crawford 325, Honolulu, HI 96822
phone: +1-808-956-3784, email: jwinter@hawaii.edu

Abstract

This chapter discusses several technical changes related to the Internet – the social semantic web and linked data, the instrumentation of natural and social processes, big data and graphing analytics, and cloud-based facial recognition – and focuses on several threats resulting from these developments. As billions, or trillions, of everyday objects, including the human body itself, are equipped with sensors, a variety of new types of data will be collected, aggregated, and linked to other personally-identifiable records. These changes transgress personal privacy boundaries and lead to unjust algorithmic discrimination and loss of anonymity, resulting in undemocratic shifts in power. Three alternative scenarios for the future Internet are presented as contrasting possibilities to explore key uncertainties about the future for the year 2045. Because the framework for the future Internet is already developed and numerous aspects of it are already appearing around us, it is essential that we critically examine these systems and associated narratives in order to stimulate meaningful discussion and design policies and systems that respect citizen concerns. By examining and testing alternative visions of the future Internet, we can more closely align its development with ethical, human-centered insight.

Keywords

data and discrimination, Internet of Things, privacy, anonymity, big data, algorithmic discrimination

Introduction

We tend to think of the Internet as something virtual that we deliberately choose to access via our computers, tablets, and smartphones. In fact, the everyday world around us, including our interactions and inferred intentions, is becoming part of the Internet, often without our realization. The ongoing instrumentation of the natural world via a variety of wireless technologies, such as radio frequency identification (RFID) and near field communication (NFC), have enabled tiny sensors and actuators to connect billions, and soon perhaps trillions, of everyday objects to the global Internet. These technical and business developments have been heralded by corporations and governments as a means to promote economic and environmental sustainability and human welfare. Research and policy discussion has focused on benefits to sectors such as logistics, transportation, energy, and the environment, with visions of enhanced disaster relief, health care, tainted food recall, farming, and environmentally-sustainable smart cities and power grids. These widespread images of the future promise greater efficiency, safety, egalitarianism, and personal convenience. However, critics have responded to this technoutopian narrative, voicing concerns about surveillance and accompanying undemocratic shifts in power, among a number of ethical and human rights concerns. Much of this is attributed to the growing consolidation of media power and the resulting influence on government regulation that has led to a restructuring of Internet standards and architecture, as well as available content. As Winseck (2003) notes, media corporations have been increasingly able to shape citizen use of the Internet. Even our identity is shaped through surveillance and control of information. This chapter examines key underlying technical changes related to the Internet, including the emergence of the social semantic web and linked data, the instrumentation of natural and social processes, big data and graphing analytics, and cloud-based facial recognition. Next, threats resulting from these developments – the erosion of privacy and merging of the public and private spheres, unjust algorithmic discrimination, and loss of anonymity – are discussed. In particular, these threats are linked to undemocratic shifts in power. Finally, three alternative scenarios for the future Internet are presented: *Galaxy of Things*, *Fractured Planet*, and *Yaoyorozu Redux*.

The changing Internet

The Internet is frequently described as promoting innovation, freedom, egalitarianism, and openness and transparency of government activities. These visions acknowledge the ethos that guided Internet development for its first few decades, drawing on the values of the original open source programmers and hackers who created the protocols enabling the Internet’s open standards, decentralization, and culture of creativity and online collaboration (e.g., Himanen, 2001). While the technical logic and early cultural shaping of the Internet was free from centralized control or commercial interests, today’s Internet operates under different rules, with power being increasingly consolidated into corporate and governmental hands due to informationalization. The “generativity” (Zittrain, 2008) that characterized the early days of the Internet has been eroding for decades, and the future Internet may move away from this ethos completely. Instead, citizens face a loss of privacy and anonymity essential for autonomy and participation in a democratic society, and unjust algorithmic discrimination threatens to exacerbate existing social and economic disparities. While technological developments do not cause social change in and of themselves (Castells, 2000, 2009), they enable it, and therefore shape our interactions and social structures. Below, several key Internet developments are discussed.

The social semantic web/linked data

The social semantic web (Berners-Lee, 2000; Breslin, Passant & Decker, 2009) is the emerging web of interlinked people and content enriched by technical standards that represent people and objects (and the links that connect them). Essentially, it is a series of machine-readable standards underlying social networking services. These developments are supported by the emergence of linked data, standards and practices for connecting structured data via the World Wide Web, creating a massive, global data space that can be navigated and processed by machine intelligence without human intervention (Heath & Bizer, 2011). This machine-to-machine (M2M) processing and “intelligence” means that, through semantic web standards, computers can increasingly understand relationships between data and perform routine tasks on our behalf. For the past decade, web data have already included a “hodgepodge of sensor data contributing, bottom-up, to machine-learning applications that gradually make more and more sense of the data that is handed to them" (O'Reilly & Battelle, 2009, p.8). Kevin Ashton, who coined the term Internet of Things (described below) in 1999, states that a goal is to empower computers

> with their own means of gathering information, so they can see, hear and smell the world for themselves, in all its random glory. RFID and sensor technology enable computers to observe, identify and understand the world – without the limitations of human-entered data. (Ashton, 2009, para. 5)

As machine-guided collection and analysis continues to grow, the scale and scope of data collection and analysis will be greatly magnified.

Instrumentation of natural and social processes

Over the past fifteen years, a growing number of sensors and actuators, including those in our mobile phones, chips embedded in our cars, smart appliances, and other common objects, have begun to blend seamlessly into our everyday environment. The Internet of Things is a paradigm encompassing a wide range of developments that enable everyday objects to be tagged and uniquely identified over the Internet (Uckelmann, Harrison & Michahelles, 2010). Although there is no single definition for the Internet of Things, competing visions agree that it relates to the integration of the physical world with the virtual world – with any object having the potential to be connected to the Internet via short-range wireless technologies, such as radio frequency identification (RFID), near field communication (NFC), or Wireless Sensor Networks (WSNs). This merging of the physical and virtual worlds is intended to increase instrumentation, tracking, and measurement of both natural and social processes:

> With so much technology and networking available at such low cost, what wouldn't you enhance? What wouldn't you connect? What information wouldn't you mine for insight? What service wouldn't you provide a customer, a citizen, a student or a patient? (IBM, 2008, para. 11)

Corporations, such as IBM ("Smarter Planet") and HP ("Central Nervous System for the Earth"), and governments, including China ("Wisdom of the Earth") and the European Union, have embraced this vision, working to develop technical standards, business practices, and policy guidelines to foster its growth. *Table 1* lists examples of Internet of Things applications that have been created or envisioned.

Logistics	e.g., supply chain management (Ashton, 2009); restocking; payment systems (Uckelmann & Harrison, 2010; Atzori, Iera & Morabito, 2010).
Health care/biomedical	e.g., ambient sensors for independent living; implantable or edible medical devices (CERP-IoT, 2010).
Environmental monitoring	e.g., natural disaster prediction, such as flood, fire, earthquake, and tsunami warning systems (CERP-IoT, 2010); chemical and gas leak identification; pollution and temperature monitoring (Hvistendahl, 2012); water potability testing.

Security	e.g., motion-sensitive camera activation; access control; radiation monitoring (Ishigaki, Matsumoto, Ichimiya & Tanaka, 2013); intrusion detection (Khan, Khan, Zaheer, & Khan, 2012).
Structural engineering	e.g., monitoring and identifying faults in buildings, roads, or bridges (Agrawal & Lal Das, 2011).
Food safety and Agriculture	e.g., testing (Hvistendahl, 2012) and recall of tainted food (CERP-IoT, 2010); monitoring hydration, chemical composition, or soil quality; livestock tracking (CERP-IoT, 2010).
Smart cities, homes, power grid	e.g., infrastructure monitoring; management of smart grids to govern cost- and resource-efficient use of energy (Khan et al., 2012; Atzori, Iera & Morabito, 2010); "Green ICT" to lower environmental impact (Vermesan et al., 2011); automatic lighting and power allocation (CERP-IoT, 2010).
Transportation	e.g., aerospace part authentication (CERP-IoT, 2010); sensor-enabled roads; assisted driving (Atzori, Iera & Morabito, 2010).

Table 1. Examples of Internet of Things applications

From the perspective of corporations and governments, the assumptions underlying this agenda are that, by networking billions or trillions of devices in the everyday environment, we can enhance business efficiency and enable continued economic growth, while making the world safer and more convenient.

Big data and graphing analytics

The amount of data flowing over the global Internet each year (e.g., Web browsing, social networking, location, and video data) is poised to pass the Zettabyte (1,000,000,000,000,000,000,000 Bytes) threshold by 2016 (Cisco, 2014). For reference, one Zettabyte is also the estimated total amount of data to have traversed the Internet since its creation in 1969. "Big data" is the term used to describe large, complex data sets that require novel data management tools. The rapid increase of real-time user data (including many novel data types) has enabled sophisticated user modeling, and there are many efforts to mine and personalize this data (Jaimes, 2010). *Big data is not just more data*. It relates also to the idea of "big graphs" that allow modeling and predicting human behaviors in their rich contexts of relationships, groups, and social influence. Governments and corporations have focused on creating sophisticated graphs of citizens' online and offline activities and aggregating these data with other sources, such as physical location, public records, and online search habits. These novel data types, coupled with enhanced data storage and analytic tools that link other personally-identifiable records, enable the construction of unique profiles. Data has become an increasingly valuable commodity, and the rationale for increased gathering and analysis is linked to the idea of endless economic growth (i.e., job production and value-added services related to big data are offered as a new frontier to stimulate economic growth).

Cloud-based facial recognition

Biometric technologies that enable one's face to be uniquely identified from a digital image, video, or in person are already part of the Internet. Sophisticated facial recognition technologies enable corporations, governments, and individuals to blend online and offline data via the convergence of social networks, data mining, and cloud computing, enabling near-instantaneous matching of subject images to online identity profiles (Keller, 2011). For example, Acquisti, Gross, and Stutzman (2011) were able to match unidentified, pseudonymous profile photos of subjects from an online dating site with their Facebook photos, as well as matching students walking around college campuses with their online records using an Internet-enabled mobile device. Related technologies are already employed by large media corporations such as Facebook, and marketers are employing them in billboards, vending machines, televisions, and home gaming systems in order to gauge viewer affect and offer customized products (Wadhwa, 2012). The Xbox One game console is accompanied by an accessory called a Kinect that uses a camera to track players' movements. The Sony PlayStation 4 also has an optional camera that performs a similar function (Ackerman, 2013). Using facial recognition, Microsoft, Sony, game companies, and their affiliates know which individuals are watching television or playing a game, as well as a wealth of other personal information, such as who one is with, what they are watching or listening to, and perhaps even what is being said (via eavesdropping and automated voice recognition). Law enforcement agencies also employ advanced facial recognition systems, and they are a core component of the United States' Next Generation Identification program (Federal Bureau of Investigation, 2014). It is expected that the sophistication and reach of these technologies will continue to grow as we move towards next-generation standards for the Web and increased data aggregation and mining. Proponents of facial recognition technologies argue that they will lead to increased security and enhance entertainment. Yet, even if one tries to avoid using the Internet, cloud-based facial recognition throughout the everyday environment enables the collection of personal data linked to a specific individual – and thus threatens both privacy and anonymity.

The erosion of privacy and collapse of the private sphere

The combination of technical developments outlined above, along with the complementary capitalization of personal data, a lack of strong regulatory intervention to protect personal data in many parts of the world, and government demands for access to citizen records as a means to prevent terrorism, have led to an erosion of personal privacy and a blurring of the public and private spheres that have underpinned Western legal discourse about privacy for centuries. A dramatic increase in *personal* data collected, stored, and transmitted, coupled with billions of devices now capable of connecting to the Internet, has led to what the European Commission, Information Society and Media (2008) refers to as a "data deluge" (p. 6). Governments and corporations, often with little or no restriction, use these data for business intelligence and consumer marketing. There is something fundamentally different-in-kind about this emerging datasphere. First, the Internet of Things relies on many tiny, often invisible, components. One does not know where or when data is being collected. Even where there are regulations requiring explicit opt-in consent, one will not know if these are being violated (Winter, 2015). Further, even if opt-in consent were enabled, the difficulty of implementing such a privacy-protecting scheme would be overwhelming (e.g., think about how to handle thousands of pop-ups at the interface level). This aspect clearly complicates regulatory or technical schemes that rely on consumer consent. Further, billions, or trillions, of everyday objects, including the human body itself, will be equipped with sensors. This opens the door for a variety of new types of data to be collected – for example, the unique communication signature of a pacemaker or insulin pump, biometric data such as one's gait or keyboard strokes, and data from sensors placed nearly anywhere that could be geared to monitor nearly anything. Finally, all of this is part of a global Internet-based system. Data will be aggregated and linked to other personally-identifiable records. Mining of big data will identify patterns that were previously not available for analysis – perhaps data that seemed innocuous or meaningless will now reveal associations we had no idea it could – and that might be harmful to us in some way (Winter, 2014). Increasingly, global flows of information will make it possible for this personal data to be accessed by a variety of sources, either legally (e.g., lax regulation) or illegally (e.g., hacking).

For some time, advertising and marketing institutions have been aggressively looking for ways to “insert themselves unfiltered into their desired customers’ domestic lives in ways that encourage consumers to accept surveillance and relationships tailored to their personal characteristics” (Turow, 2006, p. 295) via direct marketing, product placements, supermarket loyalty programs, and customized media. The Internet of Things and big data analytics will only further enhance marketers’ and advertisers’ surveillance power.

In the face of these changes, some have claimed that *privacy is dead*. For example, in 2010, Facebook founder Mark Zuckerberg stated that privacy was no longer a social norm (Johnson, 2010), a statement met with some resistance by citizens and scholars alike. Nissenbaum (2010) and boyd and Hargittai (2010), for example, highlighted the importance of understanding context when discussing privacy. In boyd and Hargittai’s study, they found that “far from being nonchalant and unconcerned about privacy matters, the majority of young adult users of Facebook are engaged with managing their privacy settings on the site at least to some extent” (para. 51).

Ongoing concern about privacy transgressions and the surveillance capabilities of the Internet have led to growing recognition of a need for technical standards and governance to “build trust and confidence in these novel technologies rather than increasing fears of total surveillance scenarios” (The European Commission, Information Society and Media, 2008, p.3). In contrast to China and the United States, the European Union has long had strict data protection regulation. These two approaches came into direct conflict in 2014, when the European Union revised its general data protection regulation, requiring lawful data processing to include explicit consent. Citizens of the European Union were also afforded the “right to be forgotten,” as well as the right to port their data to other holders (Balboni, 2012). These conflicting approaches may hinder global standards development – or data protections may be weakened and ultimately left by the wayside.

Unjust algorithmic discrimination

Data mining and profiling may lead to undesirable discrimination (Custers, 2013), as big data analytics exposes sensitive behaviors or other personal information that could be used to disadvantage certain individuals or groups by corporations or governments. For example, citizens may experience political and economic discrimination related to housing, immigration, employment, political, or health-related behaviors (Winter, 2014). What was once considered harmless chunks of information, such as your location at particular times of day, what you bought at the supermarket, what appliances are running in your home, or what individuals you spoke to or were in close proximity to at a certain time, can be used to discriminate against individuals in many ways. For example, companies might offer different services, products, or prices to individuals based on their data profile (Turow, 2006, 2012; Winter, 2014). Similarly, insurers are beginning to allocate risk differently due to big data analytics (Upturn, 2014):

> A person's future health, like their driving behavior, can also be predicted based on personal tracking to set insurance prices. At an annual conference of actuaries, consultants from Deloitte explained that they can now use thousands of "non-traditional" third party data sources, such as consumer buying history, to predict a life insurance applicant's health status with an accuracy comparable to a medical exam. (p. 6)

Insurance is designed to spread risk across a large group of people, so new forms of price differentiation will place great burdens on those with certain medical conditions (or even a data profile indicating they *might* become ill). Differentiation may also lead to increased costs for healthy individuals in low-income areas or those who drive to work at night – and both groups are disproportionately populated by vulnerable social populations (Upturn, 2014).

Even where discrimination is illegal – such as basing the approval of a mortgage based on one's race or family status – other, non-protected proxy information may be used to make the same decision to decline. Barocas and Selbst (2015) note that,

> Advocates of algorithmic techniques like data mining argue that they eliminate human biases from the decision-making process. But an algorithm is only as good as the data it works with. Data mining can inherit the prejudices of prior decision-makers or reflect the widespread biases that persist in society at large. Often, the "patterns" it discovers are simply preexisting societal patterns of inequality and exclusion. (p. 1)

As Haggerty and Ericson (2006) point out, networked surveillance allows corporations or governments to assign individuals to social groups and then monitor them, with the specific logic of that system subjecting individuals to varying levels of scrutiny. Lyon (2002) describes this differentiation as "social sorting":

> Codes, usually processed by computers, sort out transactions, interaction, visits, calls and other activities; they are invisible doors that permit access to or exclude from participation in a multitude of events, experiences, and processes. The resulting classifications are designed to influence and to manage populations and persons thus directly and indirectly affecting the choices and chances of subjects. The gates and barriers that contain, channel, and sort populations have become virtual. (Lyon, 2002, p.13)

Surveillance can also shape one's identity based on categories created by advertisers. An individual's position in this "new constellation of market segments" determines the commercial offers and communication one receives (Haggerty & Ericsson, 2006, p. 16). In many cases, algorithmic discrimination unjustly harms individuals or groups who are already socially and economically disadvantaged.

Death of anonymity

The current evolution of the Internet also threatens anonymity. There is a trend towards online identity verification, where corporations such as Google or Facebook attempt to link all online user profiles together via a "real name" policy. Unique identification allows the aggregation and mining of personal information, and users who resist may be disadvantaged by being unable to access services. As danah boyd notes, "the people who most heavily rely on pseudonyms in online spaces are those who are most marginalized by systems of power" (boyd, 2011, para. 6).

Data that have been anonymized in order to meet regulatory requirements or to quell public concern can also be "re-personalized" via data mining techniques (Schwartz & Solove, 2011). Angwin and Stecklow (2010) found that omnibus data aggregators have been exploiting technology that "matches people's real names to the pseudonyms they use on blogs, Twitter and other social networks" (para. 20). Many other anonymized large data sets have been compromised through re-identification. An early example of this was the identification of Thelma Arnold ("user number 4417749"), a would-be anonymous user of the AOL search engine. In 2006, AOL released 20 million anonymous web search queries, and journalists were quickly able to identify Arnold based on her queries, many of which revealed private aspects of her life (Barbaro & Zeller, 2006). In another case, Manfredi, Mir, Lu, and Sanchez (2014) examined the data set from the Telecom Italia Big Data Challenge, which included vehicle location and mobility data from Milan, and noted that "there is no known way to anonymize location data since spatio-temporal data is highly unique to individuals and robust to changes over extended periods of time" (p. 46). It was easy to uniquely identify drivers from just a few data points. In 2013, anonymous DNA sequences posted on Internet genealogy forums were linked to DNA donors based on publicly available data (Gymrek, McGuire, Golan, Halperin & Erlich, 2013). In each of these examples, the data were highly personal but thought to be harmless because they were "anonymous."

As awareness of corporate and governmental surveillance grows and anonymity and privacy are diminished, citizens have begun to self-censor. PEN American Center (2013), a national writers group in the United States, surveyed members and found that they engaged in self-censorship in the wake of news about mass surveillance programs run by National Security Agency that include monitoring the activities of everyday citizens. As concerns about privacy invasions and lack of anonymity mount, citizens' freedom of access to information and their ability to discuss issues relevant to democratic decision-making in their communities is limited.

Democracy and egalitarian systems

Powerful technoutopian narratives champion the Internet as a catalyst for democratic discourse and increased political participation, a platform for the emergence of the "public sphere" as envisioned by Habermas (1991). Benkler (2006), for example, sees the Internet as an online public sphere, due to the increased feedback opportunities it affords. Hindman (2009), on the other hand, has found that existing power structures have only been reinforced through media consolidation that has limited the diversity of political discussions online. The feedback-rich environment possible via the future Internet, increasingly interdependent and self-organizing, certainly has unprecedented *potential* for grass-roots political action and increased citizen involvement in governance. However, while we are promised that the Internet will enhance democracy and promote egalitarian systems, the developments noted above have more often represented undemocratic shifts in power (Winter, 2014). Given the commercial value of personal data, unethical uses of big data, and privacy concerns noted above, it is questionable whether the future Internet might enable meaningful citizen participation and governance. With technological innovation, we also need social innovation (e.g., meaningful participatory design and governance) to guide development of technical systems in order to protect ethical norms and strengthen civil society. Deliberative democratic processes which actively seek to involve members of the general population in the formation of policy are essential and require meaningful multi-stakeholder dialogue involving governments, businesses, and citizens.

Alternative Scenarios

To shift from a present-focused mindset and enable ourselves to explore, test, and evaluate alternative futures in the present, three alternative futures scenarios representing distinct possibilities for the year 2045 are outlined below: The Galaxy of Things, Fractured Planet, and Yaoyorozu Redux. Dator (2009) argues that there are four fundamental archetypes for images of the future: Continued Growth (often "Continued Economic Growth"), Collapse (due to internal or external forces), Disciplined Society (focusing on survival and fair distribution), and Transformation. In this chapter, I have combined Continued Growth and Disciplined Society into a single scenario to explore the tensions between them. These scenarios are not intended to reflect *probable* futures; rather, they present contrasting possibilities for the future Internet in the year 2045. They are designed to highlight critical concerns and opportunities related to the Internet and to help foster fruitful dialogue in the present.

The Galaxy of Things

In 2045, what we once called the Internet is now truly *everywhere*, or at least in nearly every natural or manufactured thing, including our bodies, the Mars colonies, and several automated research stations on Saturn's moons. The integration of nanotechnologies and other materials science innovations took us by surprise – after years of hearing promises about proximal future applications, we suddenly realized they were all around us. Looking back, it seemed to happen overnight. In the early 2000s, amidst fears of "terrorism," many had argued that privacy was irrelevant, selfish even. Even more, as evidence of humankind's destructive influence on the natural systems of the planet became irrefutable, over 150 nations, including the Unites States, members of the former European Union, China, Russia, and Brazil, ratified the Calcutta Protocol (named after the first large city to be to destroyed by rising sea levels) in 2027. Subsequently, any resistance faded, and we acquiesced to demands for systems such as the smart grid and homes and "green" city infrastructure that strictly measured and managed our energy consumption. As China exerted its political and economic might as the world leader of sensor networking technologies and standards development, privacy regulations were quickly relaxed. Soon, there were no restrictions on the collection or analysis of personal data by government or corporate entities. Law enforcement's encroachment on personal data was increasingly upheld by the courts. Global media enterprises continued to consolidate, and the ubiquitous deployment of near-invisible sensors led to a high degree social transparency. The "private sphere" faded from existence, a relic of history embedded in archaic laws that were no longer enforced or were entirely removed.

Sharing our data was helping the planet and seemed harmless enough at first. After all, who would really care about such minutiae as what appliances we were using or what route we took to work? What harm could come from these tiny snapshots of daily life? Looking back, it seems obvious that corporations and governments were basically the same thing, or at least working in tandem. Soon enough, biometric technologies made it near-impossible to travel, purchase items, or meet others without notice. Today, no one comes to arrest us for dissent, because speaking out is futile, and few dare risk it. Even thinking about it seems dangerous due to the predictive power of the network intelligence. We are frequently reminded that decisions based on automated systems and complex algorithms are fair, as they lack human bias. In reality, long-standing social and economic injustices seem to have increased. While many people have their most basic physical needs met, citizens are rewarded based on adherence to the “common good,” and subtle punishments are meted out to anyone who deviates. The latter receive disincentives in the form of higher prices and interest rates, reduced energy access during peak times, availability of certain jobs, and many other things. Mostly, we are safe, as long as we accept our enforced identities.

Fractured Planet

In 2045, access to the Internet is a luxury enjoyed primarily by the military elite and the super-wealthy. Even so, it’s not too reliable and hardly global. By 2018, governments faced a substantial backlash as citizens banded together to oppose oppressive surveillance regimes and policies that exacerbated the digital divide. Numerous environmental disasters brought about by human activity quickly silenced this. As Calcutta, Guangzhou, Miami, Shanghai, New York, and other coastal cities began to disappear beneath the waves, even the most recalcitrant naysayer understood that climate change was real. As regions around the world were devastated by megastorms, any resistance towards collection or analysis of our personal data was quickly silenced. After the passage of the Calcutta Protocol in 2027, corporations and governments focused on energy conservation and ecosystem monitoring via the “Intelligent Earth” strategy. This was an aggressive, Internet-of Things-enabled solution to combat global warming. In addition to deploying sensors throughout the natural and built environments, great strides were made in geoengineering and weather modification. Clouds of nanoscopic smart particles were released into the atmosphere to combat global warming. Any dissent based on concerns about possible health effects were drowned out by media sources declaring these changes to be “green” and essential for survival.

Whistleblowers soon revealed that weather modification was actually used by militaries to control the weather in battlefield conditions, and the process was also poisoning the environment by shedding toxic particles. The true goal had been the militarization of space. In addition, the “Intelligent Earth” strategy did not have the desired environmental impacts; it increased surveillance and military power while, in many cases, *increasing* energy use. As climate change continued to spin out of control, states involved with the Calcutta Protocol refused to acknowledge this policy failure, and no meaningful changes were implemented. Increasingly powerful military regimes (targeting each other) and disgruntled citizen hackers (targeting the military and other elites) engaged in cyberwarfare, effectively crippling Internet services in various regions. Since most data is stored in the cloud, these numerous security breaches and data outages led to a fractured series of smaller networks. Furthermore, amid these tensions, national powers argued over standards related to the Internet of Things, leading to additional disruptions. This halted the provision of many essentials, such as electricity, heat, food, and sanitation. Some places have fared better than others. Natural disasters, massive food shortages, and pandemics have ravaged many regions and led to nearly two-thirds of the global human population dying off. As the powerful continue to hoard resources and use their wealth to achieve whatever physical security is possible, the environment and global political and economic systems continue to degrade. We are living on borrowed time.

Yaoyorozu Redux

In 2045, the natural and virtual worlds are fully integrated, but the Internet is less noticeable due to improved material flexibility, reduced sound, and aesthetics focused on minimizing conscious impact. After ratification of the Calcutta Protocol in 2027, use of fossil fuels was severely restricted, and computing quickly took on new energy-efficient forms – organic actuators, such as stimuli-responsive gels and polymers, biological computing, and other novel forms not using conventional metal or ceramic components, became prevalent. Geoengineering initiatives to slow global warming showed early signs of success, strengthening the drive towards developing truly "green" cities that effectively reduced human impact on Earth. Sophisticated machine intelligence was embedded throughout the built environment, and smart homes and cities were able to capture energy consumption in real-time and adjust based on critical needs rule sets and past usage patterns. Existing power grids incorporated clean energy sources, such as solar and wind power. These responsive systems provided personalized feedback, advocating for specific behavioral changes; and this led to energy conservation and reduced waste. Artificial intelligences (AIs) ensured that only data that was necessary for efficient operation of these systems was collected, and individuals were able to adjust their level of desired privacy in many cases.

Reliance on AIs to monitor the environment and adjust accordingly grew, and tools such as auto-responsive flood control and pollution filtering were deemed essential to human survival. These efforts led to a corresponding decrease in military funding and divestiture from fossil fuel funds, as public pressure to invest in green systems grew upon initial success. By 2035, smart things enabled with AI were all around us, embedded in nature. Like the Shinto concept of *Yaoyorozu no Kami-gami* ("Eight Million Gods"), referring to the infinite spirits or "intelligences" present in nature, the intelligent Internet provides useful services, protecting us, helping us, and making our lives more comfortable. With a renewed emphasis on the sacredness of natural processes and artifacts, we have channeled a widespread longing to restore the natural environment and come back into alignment with nature via Internet-enabled AIs. Many believe that we could not live without them. Certainly, we could not manage the complex systems that govern our environment without their guidance.

As the first artificially intelligent entities were recognized as conscious beings, and granted legal rights, a social divide began to occur between supporters of civil rights for AIs and those who bitterly opposed them. By this point, we were so reliant on them to operate our smart environments that we gladly granted AIs oversight over many vital processes. Due to the reliance and trust most humans afforded them, AIs meeting a certain threshold were granted legal personhood. Enhanced civic participation and deliberation enabled by the sentient Internet led to more inclusive decision-making. To ensure the integrity of this system, a panel of trusted AIs was selected in 2043 as an ethics oversight committee to monitor and root out unjust algorithms and ensure transparency of government. Of course, this has threatened many people, particularly those whose power has been waning or those who oppose AI due to religious beliefs. There have been several attacks on the sentient Internet, but these have had little observable impact. While initially a small resistance, the revelation this year that AIs have added fertility inhibitors to the water supplies of overpopulated cities has led many more to speak out against them.

Conclusion

This chapter discussed several technical changes related to the Internet – the social semantic web and linked data, the instrumentation of natural and social processes, big data and graphing analytics, and cloud-based facial recognition – and described several threats resulting from these developments. The erosion of privacy, unjust algorithmic discrimination, and loss of anonymity were highlighted and linked to undemocratic shifts in power. Finally, three alternative scenarios for the future Internet were outlined as a means to explore key uncertainties about the future. As a result of the scale, complexity, and relative lack of visibility of network developments, we tend to think of them something that may occur in the future; however, technological infrastructures are in constant flux, and many of these "future" developments are already here in some form. Further, as Dourish and Bell (2011) observe, "thinking of infrastructure as stable, uniform, seamless, and universally available is clearly problematic" (pp. 28-29). Because the framework for the future Internet is already developed and numerous aspects of it are already appearing around us, it is essential that we critically examine these systems and associated narratives in order to stimulate meaningful discussion and design policies and systems that respect citizen concerns. By examining and testing alternative visions of the future Internet, we can more closely align the development of the future Internet with ethical, human-centered insight.

Citation: Winter, J. S. (2015). "Algorithmic discrimination: Big data analytics and the future of the Internet." In J. S. Winter & R. Ono (Eds.)., *The future Internet: Alternative visions* (pp.125–140). Cham: Springer.

References

Ackerman, D. (2013, November 22). "Xbox One and PlayStation 4: Facial recognition shootout." Retrieved from http://www.cnet.com/news/xbox-one-and-playstation-4-facial-recognition-shootout/

Acquisti, A., Gross, R., Stutzman, F. (2011). Faces of Facebook: Privacy in the age of augmented reality. Black Hat 2011. Retrieved from http://www.heinz.cmu.edu/~acquisti/face-recognition-study-FAQ/acquisti-faces-BLACKHAT-draft.pdf

Agrawal, S., & Lal Das, M. (2011, December). Internet of Things – a paradigm shift of future Internet applications. Paper presented at the Second International Conference on Current Trends in Technology (NUiCONE 2011), Ahmedabad.

Angwin, J., & Stecklow, S. (2010, October 12). "'Scrapers' dig deep for data on Web." *The Wall Street Journal*. Retrieved from http://online.wsj.com/article/ SB10001424052748703358504575544381288117888.html

Ashton, K. (2009, June 22). That 'Internet of Things' thing. *RFID Journal*. Retrieved from http://www.rfidjournal.com/article/view/4986

Atzori, L., Iera, A., & Morabito, G. (2010). The Internet of Things: A survey. *Computer Networks, 54*, 2787-2805.

Balboni, P. (2012). EU Commission proposal for a general data protection. *Proceedings of the 2nd Annual International Congress of u-World*. Dalian, China.

Barbaro, M., & Zeller, T. (2006). "A face is exposed for AOL searcher no. 4417749." Retrieved from http://www.nytimes.com/2006/08/09/technology/09aol.html?pagewanted=all&_r=0

Barocas, S., & Selbst, A.D. (2015). Big data's disparate impact. *California Law Review, 104*. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2477899

Benkler, Y. (2006). *The wealth of networks: How social production transforms markets and freedom*. New Haven, Ct.: Yale University Press.

Berners-Lee, T. (2000). *Weaving the Web: The past, present and future of the World Wide Web by its inventor*. London: Texere.

boyd, d., & Hargittai, E. (2010). Facebook privacy settings: Who cares? *First Monday, 15*(2). Retrieved from http://firstmonday.org/article/view/3086/2589

boyd, d. (2011, August 4). "'Real names' policies are an abuse of power." Apophenia. Retrieved from http://www.zephoria.org/thoughts/archives/2011/08/04/real-names.html

Breslin, J., Passant, A., & Decker, S. (2009). *The social semantic web*. Heidelberg: Springer-Verlag.

Castells, M. (2000). Materials for an exploratory theory of the Network Society. *British Journal of Sociology, 51*(1), 5-24.

Castells, M. (2009). *Communication power*. New York: Oxford University Press.

CERP-IoT. European Union, Cluster of European Research Projects on the Internet of Things. (2010). *Vision and challenges for realising the Internet of Things.* Brussels: European Commission – Information Society and Media.

Cisco. (2014, June 10). *The Zettabyte era: Trends and analysis*. Cisco White Paper. San Jose, Ca.: Cisco Systems. Retrieved from: http://www.cisco.com/c/en/us/solutions/collateral/service-provider/visual-networking-index-vni/VNI_Hyperconnectivity_WP.pdf

Custers, B. (2013). Data dilemmas in the information society: Introduction and overview. In B. Custers, T. Calders, B. Schermer & T. Zarsky (Eds.), *Discrimination and privacy in the information society: Data mining and profiling in large databases* (pp. 3-26). New York: Springer.

Dator, J. (2009). Alternative futures at the Manoa School. *Journal of Futures Studies*, *14*(2), 1–18.

Dourish, P., & Bell, G. (2011). *Divining a digital future: Mess and mythology in ubiquitous computing*. Cambridge, Ma.: The MIT Press.

European Commission, Information Society and Media. (2008). *Internet of Things in 2020: Roadmap for the future. European Technology Platform on Smart Systems Integration.* Version 1.1 (27 May, 2008).

Federal Bureau of Investigation. (2014, September 15). “FBI announces full operational capability of the Next Generation Identification System.” Retrieved from http://www.fbi.gov/news/pressrel/press-releases/fbi-announces-full-operational-capability-of-the-next-generation-identification-system

Gymrek, M., McGuire, A.L., Golan, D., Halperin, E., & Erlich, Y. (2013). Identifying personal genomes by surname inference. *Science, 339*(6117), 321-324.

Habermas, J. (1991). *The structural transformation of the public sphere: An inquiry into a category of Bourgeois society*. Cambridge, MA: MIT Press.

Haggerty, K.D., & Ericson, R.V. (2006). The new politics of surveillance and visibility. In K.D. Haggerty and R.V. Ericson (Eds.), *The new politics of surveillance and visibility* (pp. 3-25). Toronto: University of Toronto Press.

Heath, T., & Bizer, C. (2011). Linked data: Evolving the Web into a global data space. *Synthesis lectures on the Semantic Web: Theory and technology, 1*(1), 1-136.

Himanen, P. (2001). *The hacker ethic: A radical approach to the philosophy of business*. New York: Random House.

Hindman, M. (2009). *The myth of digital democracy*. Princeton, NJ: Princeton University Press.

Hvistendahl, M. (2012). China pushes the 'Internet of Things.' *Science*, *336*(6086), 1223-1223.

IBM. (2008). "Conversations for a smarter planet." Retrieved from: http://www.ibm.com/smarterplanet/global/files/us__en_us__general__smarterplanet_overview.pdf

Ishigaki, Y., Matsumoto, Y., Ichimiya, R., & Tanaka, K. (2013). Development of mobile radiation monitoring system utilizing smartphone and its field tests in Fukushima. *IEEE Sensors Journal, 13*(10), 3520-3526.

Jaimes, A. (2010). Data mining for user modeling and personalization in ubiquitous spaces. In H. Nakashima, H. Aghajan, & J.C. Augusto (Eds.), *Handbook of ambient intelligence and smart environments* (pp. 1015-1038). London: Springer-Verlag.

Johnson, B. (2010, 10 January). "Privacy no longer a social norm, says Facebook founder." Retrieved from http://www.theguardian.com/technology/2010/jan/11/facebook-privacy

Keller, J. (2011, September 29). "Cloud-powered facial recognition is terrifying." *The Atlantic Monthly*. Retrieved from http://www.theatlantic.com/ technology/archive/2011/09/cloud-powered-facial-recognition-is-terrifying/245867/

Khan, R., Khan, S.U., Zaheer, R., & Khan, S. (2012). Future Internet: The Internet of Things architecture, possible applications and key challenges. *10th International Conference on Frontiers of Information Technology* (pp. 257-260). DOI: 10.1109/FIT.2012.53

Lyon, D. (2002). Surveillance as social sorting: Computer codes and mobile bodies. In D. Lyon (Ed.), *Surveillance as social sorting: Privacy, risk and automated discrimination* (pp. 14-30). London, UK: Routledge.

Manfredi, N., Mir, D., Lu, S., & Sanchez, D. (2014). Differentially private models of tollgate usage: The Milan tollgate data set. *IEEE International Conference on Big Data, 2014*, 46-48.

Nissenbaum, H. (2010). *Privacy in context: Technology, policy, and the integrity of social life*. Stanford, Ca.: Stanford University Press.

O'Reilly, T., & Battelle, J. (2009). *Web squared: Web 2.0 five years on*. White paper presented at the Web 2.0 Summit, October 20-22, 2009, San Franciso, CA: O'Reilly. Retrieved from http://www.web2summit.com/web2009/public/schedule/detail/10194

PEN American Center. (2013). *Chilling effects: NSA surveillance drives U.S. writers to self-censor*. New York. PEN American Center.

Schwartz, P.M., & Solove, D. (2011). The PII problem: Privacy and a new concept of personally identifiable information. *New York University Law Review, 86*, 1814-1894.

Turow, J. (2006). Cracking the consumer code: Advertisers, anxiety and surveillance in the digital age. In K.D. Haggerty and R.V. Ericson (Eds.), *The new politics of surveillance and visibility* (pp. 279-307). Toronto: University of Toronto Press.

Turow, J. (2012). *The daily you: How the new advertising industry is defining your identity and worth*. New Haven, Ct.: Yale University Press.

Uckelmann, D., & Harrison, M. (2010). Integrated billing mechanisms in the Internet of Things to support information sharing and enable new business opportunities. *International Journal of RF Technologies: Research and Applications, 2*(2), 73-90.

Uckelmann, D., Harrison, M., & Michahelles, F. (2010). An architectural approach towards the future Internet of Things. In D. Uckelmann et al. (Eds.), *Architecting the Internet of Things* (pp.1-24). Berlin: Springer-Verlag Berlin Heidelberg.

Upturn. (2014). *Civil rights, big data, and our algorithmic future*. Retrieved from https://bigdata.fairness.io/

Vermesan, O., Friess, P., Guillemin, P., Gusmeroli, S., Sundmaeker, H., Bassi, A., Jubert, I.S., Mazura, M., Harrison, M., Eisenhauer, M., & Doody, P. (2011). Internet of Things strategic research roadmap. In O. Vermesan & P. Freiss. (Eds.), *Global technological and societal trends from smart environments and spaces to green ICT* (pp. 9-52). Aalborg: River Publishers.

Wadhwa, T. (2012, Aug 8). "What do Jell-O, Kraft, and Adidas have in common? They all want to know your face." Retrieved from http://www.forbes.com/sites /singularity/ 2012/08/08/ billboards-and-tvs-detect-your-face-and-juice-up-ads-tailored-just-for-you/

Winseck, D. (2003). Netscapes of power: Convergence, network design, walled gardens, and other strategies of control in the information age. In D. Lyon (Ed.), *Surveillance as social sorting: Privacy, risk and digital discrimination* (pp. 176-198). New York: Routledge.

Winter, J.S. (2014). Surveillance in ubiquitous network societies: Normative conflicts related to the consumer in-store supermarket experience in the context of the Internet of Things. *Ethics and Information Technology, 16*(1), 27-41. doi:10.1007/s10676-013-9332-3.

Winter, J.S. (2015). Privacy challenges for the Internet of Things. In M. Khosrow-Pour (Ed.), *Encyclopedia of Information Science and Technology*, Third edition (pp. 4373-4383). Hershey, PA: IGI Global.

Zittrain, J. (2008). *The future of the Internet – And how to stop it*. New Haven: Ct.: Yale University Press.