COGITATIO

Media and Communication (ISSN: 2183–2439)
2020, Volume 8, Issue 2, Pages 270–279
DOI: 10.17645/mac.v8i2.2761

Article

Staying at the Edge of Privacy: Edge Computing and Impersonal Extraction

Luke Munn

Institute for Culture and Society, Western Sydney University, Penrith, NSW 2751, Australia;
E-Mail: l.munn@westernsydney.edu.au

Submitted: 5 January 2020 | Accepted: 12 May 2020 | Published: 23 June 2020

Abstract
From self-driving cars to smart city sensors, billions of devices will be connected to networks in the next few years. These devices will collect vast amounts of data which needs to be processed in real-time, overwhelming centralized cloud architectures. To address this need, the industry seeks to process data closer to the source, driving a major shift from the cloud to the 'edge.' This article critically investigates the privacy implications of edge computing. It outlines the abilities introduced by the edge by drawing on two recently published scenarios, an automated license plate reader and an ethnic facial detection model. Based on these affordances, three key questions arise: what kind of data will be collected, how will this data be processed at the edge, and how will this data be 'completed' in the cloud? As a site of intermediation between user and cloud, the edge allows data to be extracted from individuals, acted on in real-time, and then abstracted or sterilized, removing identifying information before being stored in conventional data centers. The article thus argues that edge affordances establish a fundamental new 'privacy condition' while sidestepping the safeguards associated with the 'privacy proper' of personal data use. Responding effectively to these challenges will mean rethinking person-based approaches to privacy at both regulatory and citizen-led levels.

Keywords
artificial intelligence; cloud; edge computing; personal data; privacy; smart city; surveillance

Issue
This article is part of the issue "The Politics of Privacy: Communication and Media Perspectives in Privacy Research" edited by Johanna E. Möller (Johannes Gutenberg University Mainz, Germany), Jakub Nowak (Maria Curie-Skłodowska University, Poland), Sigrid Kannengießer (University of Bremen, Germany) and Judith E. Möller (University of Amsterdam, The Netherlands).

1. Introduction

Cloud architectures have reached a crisis point. From self-driving cars to smart city sensors; 30 billion devices will be connected to networks in the next few years (Stack, 2018). Yet existing cloud infrastructures are not designed for their needs. As Shi and Dustdar (2016, p. 78) explain; "the bandwidth of the networks that carry data to and from the cloud has not increased appreciably. Thus; with edge devices generating more data; the network is becoming cloud computing's bottleneck." Connected medical devices will generate huge volumes of data, connected cars will need near real-time processing, and connected cameras will capture extremely personal information. These three properties—data volume, data latency, and data privacy—are driving a shift away from the cloud model (Simonelli, 2019).

The technology industry aims to address these needs by moving computation and storage to where it is needed. Over the next few years; this will mean shifting many applications from centralized data center facilities to highly distributed devices at the edge of the network—from the center to the 'edge' or from the cloud to the 'fog.' Simply put, the edge is both a paradigm and an architecture that aims to store and process data closer to the source. Rather than having to move massive volumes of data all the way back to the cloud—a slow, expensive; or even unviable proposition—the edge processes it on site, addressing both latency and bandwidth issues. In doing so, the edge functions as a distributed

layer of intelligence deployed at a local level (Luan et al., 2015). Practically this will take the form of cameras, sensors, switches, and micro-servers installed throughout vehicles, homes, workplaces, neighborhoods, and the broader urban environment. Following Shi and Dustdar (2016, p. 79) then, an edge device is "any computing or networking resource residing between data sources and cloud-based data centers." A smartphone could act as the edge between the body and the cloud; a smart home gateway could be the edge device between the home and the cloud.

In capturing, processing, and distributing highly personal information, the shift to the edge introduces critical new challenges to privacy. Yet existing scholarship on the edge and privacy is largely constrained to computer science and security, focusing tightly on solving specific technical problems (Alrawais, Alhothaily, Hu, & Cheng, 2017; Mukherjee et al., 2017; Roman, Lopez, & Mambo, 2018; Yi, Qin, & Li, 2015). Instead, this article poses a different research question: How are privacy-related conditions modulated by the edge, and what are the social and individual implications of this modulation? If the edge was always predicted to be a technical challenge, a 'non-trivial' extension to the cloud (Bonomi, Milito, Zhu, & Addepalli, 2012) it is also a highly political technology in transforming the way data can be handled and information extracted. The article will argue that the edge allows a form of individualization without identification, shaping privacy conditions while sidestepping the harder regulatory frameworks associated with 'personal data' as it is conventionally understood.

First, this article outlines the capabilities of the edge and introduces two specific understandings of privacy. Then, it posits two edge computing scenarios: ethnic facial recognition and an automated license plate reader. In each scenario, edge devices extract data and transform it into actionable insights, but then anonymize or abstract it before transferring it back to centralized data centers. The new affordances of the edge thus introduce new decisions around data. After that, the article poses these questions: what data to collect, how to process it at the edge, and how to 'complete' it in the cloud. Finally, the article discusses the implications of this shift, in establishing an intermediate layer of intelligence between the user and the cloud, edge computing circumvents some of the traditional privacy safeguards that have focused heavily on personal data collection and cloud storage.

2. Privacy Proper vs. Privacy Conditions

Much of the computer science literature surrounding privacy and the edge has focused on the security of personal data. While cloud data centers have developed a formidable array of hardware and software security features over time, edge-based hardware—consumer products like cameras, phones, and wearable devices—often only feature consumer-level protections. Such devices are often 'resource poor,' their micro-controllers were not designed for connectivity and lack the processing power to run cryptographic procedures, resulting in security issues such as authentication, access control, and data protection (Alrawais et al., 2017, p. 35). Moreover, rather than the closed ecosystem of the centralized data center, where a company can control access to servers, edge networks are a far more open, unrestricted architecture composed of potentially hundreds or thousands of devices, often operated by different providers. Because of this lack of a global perimeter, edge computing is more susceptible to rogue gateway attacks, where network nodes pretend to be legitimate and coax users to share data with them (Roman et al., 2018, p. 13). Edge hardware thus presents a highly vulnerable site, open to exploits. Stored on a diverse array of devices, many with minimal consumer-grade protections, this information presents a rich target for leaks and hacks. Already there have been a number of high-profile attacks exploiting these weaknesses. From cardiac devices at St Jude's Hospital to the TRENDnet webcam hack and the Mirai botnet that caused large sections of the Internet to go down, these examples demonstrate that "much of the embedded firmware running connected devices is insecure and highly vulnerable, leaving an indeterminate number of critical systems at risk" (Dunlap, 2017).

If edge hardware itself is more vulnerable, the content it captures only amplifies these concerns. Edge devices have significantly more potential to collect highly personal and highly detailed data. From health monitors to home assistants, many of these devices will be physically close to users or situated in the heart of their living environments, capturing more intimate data. The onboard camera of a self-driving car, for instance, will be switched on and recording for the entire duration of a driving session (Bloom, Tan, Ramjohn, & Bauer, 2017). It might be capturing the driver's face, but also her surrounding world, including her children and passengers. Edge-computing means this data no longer has to be heavily compressed snapshots that can be transferred to the cloud. From a privacy perspective, lifting this technical constraint means that a camera can be both higher resolution and lower latency, allowing the capture of a glance, for instance, at 60 frames per second. Moreover, devices on the edge, whether comprising a smartphone, a wearable device, a vehicle, or a network of cameras distributed throughout an urban space, have the potential to capture fine-grained location data. Locational data alone is highly valuable, aggregated over time it becomes a timeline of an individual's movements, providing an incisive window into their habits, behaviors, and preferences. As Barreneche and Wilken (2015) assert, such locational data becomes a sophisticated form of 'geodemographic profiling' that can then be leveraged for predictive purposes. Already we've seen how such locational data can be used to harass and target individuals, whether by law enforcement agencies (Munn, 2018) or private companies (Hill, 2014).

Overall, then, edge literature conflates privacy with personal data security. In this view, while the problem is technically challenging, it is theoretically straightforward—translating existing technologies like encryption to the edge will ensure 'privacy' for all (Zhang, Chen, Zhao, Cheng, & Hu, 2018; Zhang, Wang, Du, & Guizani, 2018).

In focusing on the security of personal data, edge literature coheres closely to the concept of privacy developed by the General Data Projection Regulation (GDPR), a version of privacy I will term 'privacy proper.' While the GDPR was conceived in the European Union, its definitions and framings have been taken up by various countries around the world. Applied to over 500 million citizens, the GDPR now forms one of the "de facto global standards for data privacy and protection" (Barrett, 2019). For the GDPR, personal data is key. "The term 'personal data' is the entryway" to the application of the regulation, states the law, "only if a processing of data concerns personal data" does the GDPR apply (European Commission, 2018). What exactly constitutes personal data? The regulation states that "data must therefore be assignable to identified or identifiable living persons to be considered personal" (European Commission, 2018). Once data conforms to this definition, the company or agency becomes a 'data processor' who must maintain compliance—hamstrung in terms of what kinds of data may be captured, how it may be stored and accessed, and which borders it may cross. In this article, privacy proper will thus designate a threshold that actors do not wish to cross, a regulatory minefield triggered when organizations begin dealing with personal data.

Certainly, the edge presents some obvious challenges for personal data security. Yet more subtly, the edge may shape privacy-related abilities without necessarily processing or storing personal data. To differentiate this possibility, I introduce a second term, 'privacy condition,' drawing on recent work by legal scholar Julie Cohen. To rescue privacy from vague rhetoric and unenforceable ideals, Cohen begins not from the figure of the self, but from the ground of the underlying conditions "that are needed to produce sufficiently private and privacy-valuing subjects" (Cohen, 2019, p. 1). If rights discourse and legal rhetoric can be abstract, then conditions have a specificity, a concreteness. Conditions actively enable some privacy-related abilities while making others improbable or even impossible. Digital data and the sharing of information has made the stakes of these abilities and inabilities suddenly very clear. This is why, even though privacy clearly has a long lineage in liberal political philosophy, Cohen stresses that privacy is a "paradigmatic information-era right," one not defined by rights discourse, but by the conditions established within the "political economy of informationalism" (Cohen, 2019, p. 2; see also Cohen, 2017).

Given this framing, Cohen wants to focus on the particular set of "design, production, and operational practices' most likely to produce privacy-valuing conditions" (Cohen, 2019, p. 1). A distinct version of privacy emerges from a set of affordances, the possible range of uses made available by an object or environment (Cohen, 2019, p. 12). In other words, particular privacy conditions emerge from particular technical configurations. As a nascent technology, the edge enables new affordances, allowing subjects to be apprehended, mediated, and responded to in distinct new ways—even when, or especially when, so-called personal data is never handled. With these two terms defined, the following scenarios focus on how edge affordances modulate privacy conditions while allowing actors to sidestep the requirements attached to privacy proper.

3. Two Scenarios

The first scenario is license plate capture and analysis, drawn from a recent article on hybrid cloud-edge computing (Zhang, Zhang, Shi, & Zhong, 2018). The authors lament the siloed nature of current data collection. They suggest that many public and private agencies would have an interest in obtaining license plate data in order to understand where citizens are located and where they travel to. Yet due to anxieties around data sharing and user authentication, each of these institutions conducts their own data capture and maintains their own cloud-based repository. The result is that "data owned by multiple stakeholders is rarely shared among data owners" (Zhang et al., 2018, p. 2004).

The edge introduces new possibilities into this scenario. For one, edge nodes can act as a nexus, combining data sources from multiple stakeholders. As Zhang et al. (2018, p. 2005) suggest, footage from the on-body cameras of police officers could be combined with squad car camera feeds, mobile uploads and more traditional CCTV feeds to form a far more extensive and comprehensive data source. Data can be assembled from various sources, processed in order to remove sensitive information, and then distributed to stakeholders. The edge's ability to decouple data collection from data storage thus has the potential to foster formerly unworkable alliances. For instance, Zhang et al. (2018, p. 2005) note that both private insurance companies and public district health boards would be interested in some of the same data. Groups of institutions might band together to collect license plate data, smart city data, or health data, creating broad infrastructures of surveillance.

Of course, this indiscriminate surveillance introduces a range of problems if privacy proper is invoked. Any one of these raw video feeds might capture facial details that could be used to identify an individual, overstepping the privacy boundaries allowed by an institution. Yet the edge can again provide a solution. In aggregating this data at a site long before it arrives back at the cloud, the edge acts as a kind of pre-processor for data. Rather than transferring all of the raw video data back to the cloud, Zhang et al. (2018, p. 2005) propose that the edge node conducts video clipping, scanning "video streams to se-

lectively filter out frames with a license plate." Edge computation would locate only those frames where black letters on a white background indicate a license plate. Each frame is then cropped to only show the plate, and optical character recognition technology converts the plate photograph to its alphanumeric equivalent, e.g., CLA974. Finally, this small text field is transferred to the cloud facilities of each stakeholder. By introducing an intermediary layer between capture and cloud, between user and stakeholder, the edge also introduces a new set of privacy-challenging affordances. Data can be collected from multiple stakeholders but then parsed at the edge, selected, sampled, and scrubbed before continuing on to cloud-based facilities. These dynamics can also be seen in the next scenario of edge computing.

The second scenario is the facial detection of ethnic minorities. In 2019, Wang, Zhang, Liu, Liu, and Miao published the article 'Facial Feature Discovery for Ethnicity Recognition.' While this article was not explicitly posited as an edge application, speculating about its transfer to this domain is hardly a leap. Indeed, as a slew of recent technical articles suggest, researchers are already embracing the new possibilities that edge computing offers for facial detection in urban areas (Dautov et al., 2018), crowd monitoring (Bailas, Marsden, Zhang, O'Connor, & Little, 2018), and intelligent surveillance (Hu et al., 2018), with one going so far as to call real-time video analytics the edge's 'killer app' (Ananthanarayanan et al., 2017).

Wang, Zhang, and Taleb (2018, p. 1) begin by noting that "the analysis of race, nation, and ethnical groups based on facial images is a popular topic recently in face recognition community." Bypassing even the barest consideration of ethics, the authors suggest that this new field would naturally be beneficial for state actors wishing to enforce certain restrictions on their citizens: "With rapid advance of people globalization…face recognition has great application potential in border control, customs check, and public security" (Wang et al., 2018, p. 1) The disturbing enthusiasm for such privacy-impinging surveillance is not limited to China, but is increasingly evident across cities in Ecuador, Pakistan, Kenya, Germany, and the United Arab Emirates (Mozur, Kessel, & Chan, 2019).

Yet, frustratingly for the article's authors, ethnicity can often be difficult to detect, either because the morphologies of race are too subtle or because the individual contains traces of multiple ethnicities. The problem, from an engineering perspective, is that "the gene of one ethnical group is hardly unique and it may include various gene fragments from some other ethnical groups" (Wang et al., 2018). Fortunately, facial aspects can be analyzed in a far more fine-grained manner through computational technologies in order to reveal their ethnicities. The authors set about identifying three ethnic groups: Uighur, Tibetan, and Korean (Wang et al., 2018). The article, like many in machine learning, essentially lays out the steps used to produce the model and measures its effectiveness against competing models. The model is trained on an image set of university students, and gradually learns to identify the three ethnic groups with more success, displaying progressively lower levels of uncertainty.

Key for the authors' model is the extraction of a 'T' feature from the center of each photograph containing the lips and nose (Wang et al., 2018). While the T varies with each ethnic group, these morphological features are considered to be the telltale markings that distinguish whether an individual is within the targeted ethnic group. Indeed, the extraction of the T, while obviously deleting key facial information, amplifies the model's ability to detect ethnicity. As the authors note that "actually, the facial features extracted from the 'T' regions are more suitable for ethnicity recognition since the unrelated information has been filtered out" (Wang et al., 2018). In this application, the full photograph of the individual is unnecessary or even a distraction. The model does not need to do the computationally intensive work of facial identification—who exactly an individual is—but rather the simpler task of determining whether an individual is 'ethnic' or not.

Such a technology would seem tailor made for the edge. As more cameras are connected to networks, the possibilities of surveillance grow. However, video data itself is massive, becoming both economically expensive and technically infeasible if it is sent back to the cloud. As the authors of one study suggested, processing raw video from widely distributed "CCTV cameras and mobile cameras not only incurs uncertainty in data transfer and timing but also poses significant overhead and delay to the communication networks" (Nikouei et al., 2018, p. 1). In the cloud model, images need to be sent from all the cameras to a data center facility via the network, be processed in this centralized facility, and then the result delivered to a client or end-user. This lengthy process not only introduces significant latency, but makes some surveillance applications essentially unviable from a technical perspective.

Instead, the edge allows processing to be conducted at the source. No identifying image needs to be sent back to the cloud and compared against an exhaustive database of citizens. No personal data is 'collected' by the agency in the sense of being transmitted to a data center where it will be held indefinitely in a database or stored on a hard disk. Instead, this machine-learned model could be compressed and loaded onto a small edge-based device with a camera. Such a device would then process its image feed in real-time, rapidly determining whether an individual is 'ethnic' or not.

Once determined, this compressed yet highly consequential piece of information might be used in any number of ways. In border security, for instance, one could imagine a green light turning red and a passenger selected for additional screening. In a smart city scenario, this data might be paired with a camera's location and uploaded to form an aggregated portrait of ethnic populations over time. Such data is based on an individual but rendered impersonal, providing insights for gover-

nance while sidestepping the harder restrictions around personal data. From a broader political perspective, the scenario demonstrates how edge affordances might underpin new forms of less governed control, establishing a privacy condition that avoids directly confronting the regulatory apparatus attached to privacy proper.

4. Questions for Privacy at the Edge

The edge complicates established privacy conditions, reopening critical debates about the ways such informational architectures may impact the everyday lives of individuals and amplify existing power asymmetries. While the scenarios above raised some of these issues indirectly, in the next sections they form three explicit questions.

4.1. What Data Will Be Collected at the Edge?

As millions of new devices are connected to networks over the next few years, the possibilities of data capture will also proliferate. As discussed, these devices, located in the home, on the wrist, or stationed around the neighborhood, will be able to capture fine-grained, highly personal data. While some network constraints will certainly still persist, edge computing means that data collection practices are no longer dictated so tightly by transmission back to the cloud. Indeed, as mentioned above, it is precisely these possibilities that have led to the many articles on real-time monitoring, crowd monitoring and 'intelligent surveillance' via edge computing (Ananthanarayanan et al., 2017; Bailas et al., 2018; Dautov et al., 2018; Hu et al., 2018). This literature displays a general rush to embrace these possibilities, even though these applications have clear implications for privacy intrusions and personal freedoms. What these enthusiastic responses demonstrate is that in many ways it was technics, rather than ethics, which limited the extent of previous intrusions into personal data. Network speeds, bandwidth capacities and physical distance were hard restrictions. To a significant degree, edge computing lifts these constraints, providing more freedom to public and private actors wishing to delve further into individuals and their lives.

These capabilities mean that the question of data collection will hinge less on technical and economic concerns (cost to transfer gigabytes back to the cloud) and more on company culture, ethical values, and policy stipulations—if these are even in place. With technical constraints lifted, companies will be under increased pressure to collect more, and more intrusive data, which could provide key business insights. Yet individual companies are not entirely free in navigating this ethical terrain. Companies do not operate in isolation, but within competitive industries, particularly the highly contested technology field. Given these conditions, companies are subject to the "coercion of competition" (Marx, 2004, p. 675). If one company chooses not to push the ethical boundaries of data capture, others will (Kokalitcheva, 2019). At a time when comprehensive data has become highly valuable, this decision grants one company strategic advantage over their competitors.

4.2. How Will Data Be Processed at the Edge?

The edge introduces an additional layer of mediation between users and the cloud, forming a site for processing data after it has been captured, but before it is stored and centralized. As suggested by the scenarios above, this interposition creates new possibilities for data processing. Rich, highly detailed data can be captured by edge devices and then processed by an edge hub in order to extract nuggets of valuable information, which is then passed back to a centralized cloud facility.

Abstraction becomes a key term within this process. How will highly personal data be transformed into impersonal, anonymized data? Here edge computing can draw on a number of existing technologies, from k-anonymity (Sweeney, 2002) to micro-aggregation (Domingo-Ferrer, Sánchez, & Soria-Comas, 2016). These established techniques, broadly applicable to any information set, include substitution, in which identifying values are randomly replaced, shuffling, so that associations between variables are lost, sampling, in which a partial set from the whole is transmitted, and variance, in which numerical values are perturbed or altered (Curzon, Almehmadi, & El-Khatib, 2019). Certainly, such technologies provide established means of handling particular types of data and aspects of applications. Yet they can also become a way of black-boxing problems and arriving too quickly at a 'privacy solution.'

Instead, the task is to keep the question of data extraction in the foreground: How is data mediated at the edge and what is lost or gained in this intermediation? Highly specific location data, for instance, might be captured at the edge, but then generalized into a district or combined with other user locations. A gender field might be used in an edge calculation, but then dropped, something users may or may not want. An individual's race might be clumped into a parent category, imposing a statistical system and erasing specific origins. In every permutation, a slightly different data subject is rendered (see Cheney-Lippold, 2018; Koopman, 2019). These examples stress that the technical transformation of information also has political and social implications. Abstraction, then, should be seen less as a solution and more as a set of design decisions around data. These decisions come together to form a particular configuration of practices and protocols, establishing a privacy condition that imposes itself on subjects in certain ways.

For those tasked with making these design decisions, abstraction attempts to walk a tightrope, balancing the desire of states and corporations to "capture it all" against the desire of individuals and their "right to be let alone" (Warren & Brandeis, 1890, p. 193). To claim that nothing should ever be captured would be naïve, to

claim that everything should would be unethical: "There is a natural tension between the quality of data and the techniques that provide anonymity protection," observes Latanya Sweeney (2001, p. 33); given these tensions, the goal is to design an optimal release so that "the data remain practically useful yet rendered minimally invasive to privacy." For both public and private actors, capturing valuable data while remaining sensitive to privacy issues will take care and consideration.

4.3. How Is Data Completed in the Cloud?

If the edge is a site of intermediation, the cloud becomes the site of completion, where data is assembled together, integrated into more formal structures, and processed for additional insights. Completion stresses the aim of both public governments and private corporations to exhaustively analyze data. If data is capital, then in order to accrue more value, one must extract more data from more subjects, accumulate it in increasingly larger volumes, and mine it incessantly for insights (Sadowski, 2019). Here the resource-constrained environment of the edge leans heavily on the resource-rich environment of the cloud. Indeed, new data center architectures embrace this role as a site of intensive processing, developing dedicated chips with liquid cooling in order to support the heavy computation required by machine learning applications (Sverdlik, 2018). If these conditions are highly technical, the insights they derive from high intensity processing shapes privacy conditions in concrete ways.

Completion foregrounds the design of a data pipeline. Decisions made about (1) what data to capture and (2) how the edge processes that data must also take into account (3) how the cloud processes that data to arrive at productive insights. Here the cloud and the edge might supplement each other with their respective strengths and weaknesses. The edge is decentralized, with low latency but low power, capable of capturing much but processing, storing, and transmitting little. The cloud is centralized, ill-suited for capture with its high latency but excellent at processing and storage. Given these trade-offs, the edge needs to deliver low volume but high potential data that can be intensively processed by the cloud to generate value.

As the two scenarios discussed above suggested, a circuit for completing data and maximizing its value is already emerging. First, data is collected from devices distributed at the edge. This data is then distributed to the closest edge node, processed in order to clean up or sample the data, and then passed onto a centralized cloud facility, where it is assembled into a training set of machine learning. High intensity processing in the cloud is used to train a model based on this dataset, gradually becoming better over time. Once completed, the machine learning model is then compressed into a light-weight version and distributed back out to edge devices, where it can function autonomously.

Here we see a feedback loop, where captured information becomes training data, which in turn contributes to more comprehensive mechanisms of capture. Indeed, whole companies have emerged based on riding this loop of "embedding edge intelligence as close to the source of streaming sensor data as possible" (Foghorn Systems, 2019). If the realization of this approach is still nascent, it is clear that developing such machinic intelligence will follow the blueprint already laid down by broader regimes of technical capture and data analysis. The imperative is to more fully apprehend the individual and her lifeworld, to more exhaustively grasp her properties, her practices, and her sociocultural milieu (Harcourt, 2015; Pasquinelli, 2015; Steyerl, 2016). This circuit thus strives to delve ever further into the subject and her everyday life, gradually apprehending her bodily characteristics, daily behaviors, location over time, and social affiliations (Finn, Wright, & Friedewald, 2013), until no secrets remain.

While a company may complete its own data, completion might also be undertaken in a more unauthorized or unexpected way by others. Data collected at the edge might well be anonymised in a robust way before transmission to cloud storage facilities. However, as scholars have shown, data can be de-anonymised by integrating multiple datasets together and then cross-indexing values against each other (Narayanan & Shmatikov, 2008; Sweeney, 1997). Promises of unassailable privacy are often broken promises (Ohm, 2010). If the edge introduces a new set of decisions about how data will be appropriately handled and transformed, these scenarios warn companies and organizations that they must also take into account the combinatorial possibilities of the cloud as well.

5. End Run Around Privacy Protections

If edge computing holds out enticing promises, its abilities may also impinge on the freedoms of individuals and the rights of communities. In this sense, edge computing forms the latest incarnation of what Shoshana Zuboff (2019) has described as surveillance capitalism. For Zuboff (2015, p. 83), surveillance capitalism accumulates "not only surveillance assets and capital, but also rights" through "processes that operate outside the auspices of legitimate democratic mechanism." Yet counter to Zuboff, rather than acquiring rights, these technical processes seek to never invoke rights. If big data accomplishes an "end run around procedural privacy protections" (Barocas & Nissenbaum, 2014, p. 31), then edge computing also carries out an 'end run' of its own. The goal is to extract data, value, and capital while never venturing into the legal and ethical minefield of privacy proper.

One way of doing this is to respond to the individual while filtering out, deleting, or abstracting away data deemed to be personal. The small, hyperlocal devices of the edge, situated in a smart home or a smart

city, will be far more adept at latching onto the behaviors and bodies of individuals. The edge can respond to these inputs in the moment, without storing the names and identifiers typically associated with 'personal data.' As a site of preprocessing, the edge is able to draw upon single bodies and personal lives, yet immediately abstract this data or aggregate it into a depersonalized mass. In this sense, the edge resonates with Antoinette Rouvroy's observation that algorithmic governance strives to never confront the person in her entirety, to never directly call her up as a political subject. 'The only subject' such governmentality needs, Rouvroy (2013, p. 154) stresses, is a "unique, supra-individual, constantly reconfigurated 'statistical body' made of the infra-individual digital traces of impersonal, disparate, heterogeneous, dividualized facets of daily life and interactions." A subject is apprehended at an individual level, but not necessarily identified.

Indeed, running through all these edge scenarios is the sense that the former key question—whether or not a user can be identified—may be subsumed by a far more fundamental question: What forms of life are being extracted from the user *even though* they are not identified? The de facto framing of privacy proper ushered in by the GDPR has privileged personal data. Yet this entire legal edifice of protections only applies once this definition is reached. Perhaps data never needed to be personal to be valuable. Perhaps control may be enacted and maintained without identifying a unique individual. Indeed, recent work on group privacy (Floridi, 2014; Mittelstadt, 2017; Taylor, Floridi, & Van der Sloot, 2016) responds precisely to this realization. Even without explicit identification, the new spaces enabled by edge computing present a verdant territory for extractive regimes (Mezzadra & Neilson, 2019), a rich zone of markers and moments to capture and respond to. While this extractivist logic deals with each person in turn—capturing moods and faces, responding to bodies and individual inputs, identifying movements and work performances—its value is only obtained by aggregating this data, by assembling and mining it en masse. This is why Tiziana Terranova (2018, p. 1) stresses that the "extractivism of data capital" siphons off the energetic behaviors and activities of the broader social body. The edge suggests a form of extraction that is individualized but not personalized.

If the individual-but-impersonal is one way of carrying out an end run around privacy, then another is avoiding some of the sharper points of personal data laws. While several existing laws regulate data that is 'held' (Mexican Congress, 2010) or 'stored' (U.S. Congress, 2018), the edge provides a new intermediary layer of intelligence where data can be captured, derived from, and then discarded or fundamentally transformed before it is stored. Through this affordance, the edge establishes a new frontier site for processing, a grey zone that seems sparsely covered by existing legislation, which has so far focused heavily on a centralized cloud model. The technology industry is all-too-aware of this possibility, even if it is framed as law abiding. "To avoid breaking the new law and thus being fined, companies should keep most of the data collected out of the cloud and process it at the edge" recommends one tech pundit (Valerio, 2018). Far more effective than eroding privacy is never confronting privacy proper to begin with.

How might we respond to the new privacy conditions instantiated by edge architectures? Regulators and policymakers will need to develop a broader and more nuanced understanding of the cloud. If centralized, hyperscale data centers remain at the core of cloud computation, the edge connects a cascading set of devices from regional hubs all the way down to local base stations and wearable and personal devices. These devices, though low-powered and often overlooked, form the new frontier for data collection practices, passing information streams up the chain, where it is aggregated together before finally arriving at the traditional data center. Yet if this ecosystem is vast, it is not monolithic. Devices at each level have distinct capacities. For example, the heavy encryption assumed in a full-scale data center may be impossible on many low-end edge devices. Regulation will thus need to be expansive but also articulated, developing codes and guidelines appropriate for each level of this architecture.

Along with acknowledging the new constellation of architectures that comprise the cloud, regulation also needs to address the edge's more situated, responsive capabilities. As the scenarios above suggested, the concept of storing an individual's personal information in a centralized database comes at the end of a long chain of activities and possibilities—or never at all. While edge devices certainly function as key points of capture, they will also carry out important processing operations, especially as hardware and software within this nascent field matures. Machine learning models, as discussed, are already being embedded in edge devices, meaning that facial detection, video trimming, and other key operations can take place within the device itself in real-time. Such data may be retained, abstracted into less 'personal' forms and transmitted back to the cloud—or simply discarded to make way for the next interaction. In doing so, edge-based devices will allow individualized interactions in the moment without having to fully confront the person and her associated rights. These technical abilities thus require a political and epistemological shift in privacy safeguards. Rather than beginning from the autonomous individual and her bundle of rights, rethinking privacy conditions from an operational standpoint as Cohen did might prove more suited for our era of rapid technological change.

For their part, citizens, activists, and organizations might productively question this model of personal privacy. Instead of this person-based approach, they might move to more communal models, based on the group, the neighborhood, the city, or the broader community. Evgeny Morozov (2015, 2018a, 2018b) has been at the

forefront of this questioning, long arguing that the current model provides nominal protection for the individual, while continuing to funnel valuable data to tech giants—Google, Amazon, Facebook, and others—who monetize it for financial gain. Instead, he suggests a socialized infrastructure where citizens could pool together their data. This public data commons can be leveraged by technologies for the public good, directing value back into the hands of the data producers. This approach recognizes that individuals have little purchase on a political economy predicated on de-individualized, aggregated data. Instead, thinkers like Morozov and other data commons advocates (Jarman & Luna-Reyes, 2016; Shkabatur, 2018; Simon, 2018) join group privacy theorists (Mittelstadt, 2017; Taylor et al., 2016) in recognizing that privacy demands are both politically amplified and technically clarified when coming from a community. What would this community-based understanding of privacy look like on an everyday operational level? Due to the edge's emergent nature, more work is needed to bring together the technical, social, and legal and develop a workable privacy model attentive to the novel conditions that the edge introduces.

6. Conclusion

This article has explored how the shift to edge computing introduces new privacy challenges. While widely covered in engineering and computer science research, there have been few, if any, studies on the cultural, social, and political implications of edge computing. Given this gap, this article has merely introduced some key concepts and sketched out some initial possibilities. More research is urgently needed to examine the tensions and decisions ushered in by this paradigm, moving beyond technical capabilities to focus on social and ethical responsibilities. After defining the edge and two framings of privacy, the article posited two scenarios drawn from real-world engineering articles: an ethnic facial detection model and an automated license plate reader. While personal data security has been the traditional focus, these scenarios suggested that the edge poses a more subtle and significant set of questions. The technical affordances of the edge allow data to be captured, processed, and completed in new ways. Such decisions establish a significant privacy condition, shaping the ways in which consumers are targeted and the methods by which subjects are governed. They suggest that asymmetric power relations might be amplified while avoiding existing privacy regulation, slipping through the definitions of personal data established by current data safeguards. In this sense, novel network architectures open up a legal and ethical loophole. If the edge is seen as a technical solution, it also presents a political solution, facilitating a mode of power able to target the individual without crossing the threshold of privacy proper.

Acknowledgments

This research work was supported as part of the wider project 'Data Centres and the Governance of Labour and Territory,' an Australian Research Council Discovery Project (DP160103307).

Conflict of Interests

The author declares no conflict of interests.

References

Alrawais, A., Alhothaily, A., Hu, C., & Cheng, X. (2017). Fog computing for the Internet of Things: Security and privacy issues. *IEEE Internet Computing*, *21*(2), 34–42.

Ananthanarayanan, G., Bahl, P., Bodík, P., Chintalapudi, K., Philipose, M., Ravindranath, L., & Sinha, S. (2017). Real-time video analytics: The killer app for edge computing. *Computer*, *50*(10), 58–67.

Bailas, C., Marsden, M., Zhang, D., O'Connor, N. E., & Little, S. (2018). Performance of video processing at the edge for crowd-monitoring applications. In H. Mueller, Y. Rongshan, & A. Skarmeta (Eds.), *2018 IEEE 4th world forum on Internet of Things (WF-IoT)* (pp. 482–487). Washington, DC: IEEE Computer Society.

Barocas, S., & Nissenbaum, H. (2014). Big data's end run around procedural privacy protections. *Communications of the ACM*, *57*(11), 31–33.

Barreneche, C., & Wilken, R. (2015). Platform specificity and the politics of location data extraction. *European Journal of Cultural Studies*, *18*(4/5), 497–513.

Barrett, C. (2019). Are the EU GDPR and the California CCPA becoming the de facto global standards for data privacy and protection? *Scitech Lawyer, Chicago*, *15*(3), 24–29.

Bloom, C., Tan, J., Ramjohn, J., & Bauer, L. (2017). *Self-driving cars and data collection: Privacy perceptions of networked autonomous vehicles*. Paper presented at the 13th Symposium on Usable Privacy and Security (SOUPS 2017), Santa Clara, CA.

Bonomi, F., Milito, R., Zhu, J., & Addepalli, S. (2012). Fog computing and its role in the internet of things. In M. Gerla & D. Huang (Eds.), *Proceedings of the first edition of the MCC workshop on Mobile cloud computing: MCC '12* (pp. 13–16). New Yor, NY: ACM Press.

Cheney-Lippold, J. (2018). *We are data: Algorithms and the making of our digital selves*. New York, NY: NYU Press.

Cohen, J. E. (2017). Affording fundamental rights: A provocation inspired by Mireille Hildebrandt. *Critical Analysis of Law*, *4*(1), 78-90.

Cohen, J. E. (2019). Turning privacy inside out. *Theoretical Inquiries in Law*, *20*(1), 1–32.

Curzon, J., Almehmadi, A., & El-Khatib, K. (2019). A survey of privacy enhancing technologies for smart cities. *Pervasive and Mobile Computing*, *55*, 76–95.

Dautov, R., Distefano, S., Bruneo, D., Longo, F., Merlino, G., Puliafito, A., & Buyya, R. (2018). Metropolitan intelligent surveillance systems for urban areas by harnessing IoT and edge computing paradigms. *Software: Practice and experience*, *48*(8), 1475–1492.

Domingo-Ferrer, J., Sánchez, D., & Soria-Comas, J. (2016). Database anonymization: Privacy models, data utility, and microaggregation-based inter-model connections. *Synthesis Lectures on Information Security, Privacy, & Trust*, *8*(1), 1–136.

Dunlap, T. (2017, May 10). The 5 worst examples of IoT hacking and vulnerabilities in recorded history. *IoT For All*. Retrieved from https://www.iotforall.com/5-worst-iot-hacking-vulnerabilities

European Commission. (2018). *General Data Protection Regulation* (2016/679). Brussels: European Commission.

Finn, R. L., Wright, D., & Friedewald, M. (2013). Seven types of privacy. In S. Gutwirth, R. Leenes, P. de Hert, & Y. Poullet (Eds.), *European data protection: Coming of age* (pp. 3–32). Dordrecht: Springer. https://doi.org/10.1007/978-94-007-5170-5_1

Floridi, L. (2014). Open data, data protection, and group privacy. *Philosophy & Technology*, *27*(1), 1–3.

FogHorn Systems. (2019, July 1). FogHorn Lightning. *FogHorn Systems*. Retrieved from https://www.foghorn.io/lightning-iot-edge-computing

Harcourt, B. (2015). *Exposed: Desire and disobedience in the digital age*. Cambridge, MA: Harvard University Press.

Hill, K. (2014, October 3). "God view": Uber allegedly stalked users for party-goers' viewing pleasure (updated). *Forbes*. Retrieved from https://www.forbes.com/sites/kashmirhill/2014/10/03/god-view-uber-allegedly-stalked-users-for-party-goers-viewing-pleasure

Hu, H., Shan, H., Zheng, Z., Huang, Z., Cai, C., Wang, C., . . . Quek, T. Q. S. (2018). Intelligent video surveillance based on mobile edge networks. In S. Li (Ed.), *2018 IEEE international conference on communication systems (ICCS)* (pp. 286–291). New York, NY: IEEE. https://doi.org/10.1109/ICCS.2018.8689194

Jarman, H., & Luna-Reyes, L. F. (2016). *Private data and public value: Governance, green consumption, and sustainable supply chains*. New York, NY: Springer.

Kokalitcheva, K. (2019, July 2). A wave of tech companies have pushed ethical boundaries to maximize profit. *Axios*. Retrieved from https://www.axios.com/big-tech-companies-ethics-facebook-doordash-59abf670-078d-4e15-9033-c0e6e6d75127.html

Koopman, C. (2019). *How we became our data: A genealogy of the informational person*. Chicago, IL: University of Chicago Press.

Luan, T. H., Gao, L., Li, Z., Xiang, Y., Wei, G., & Sun, L. (2015). Fog computing: Focusing on mobile users at the edge. *Cornell University.* Retrieved from http://arxiv.org/abs/1502.01815

Marx, K. (2004). *Capital: A critique of political economy* (B. Fowkes, Transl.). London: Penguin Books.

Mexican Congress. (2010). *Federal law on protection of personal data held by individuals* (DOF 05-07-2010). Mexico City: Mexican Congress.

Mezzadra, S., & Neilson, B. (2019). *The politics of operations: Excavating contemporary capitalism*. Durham, NC: Duke University Press.

Mittelstadt, B. (2017). From individual to group privacy in big data analytics. *Philosophy & Technology*, *30*(4), 475–494.

Morozov, E. (2015). Socialize the data centres! *New Left Review*, *91*(1), 45–66.

Morozov, E. (2018a, January). *DECODE presents: Data commons and the city* [Video file]. Retrieved from https://www.youtube.com/watch?v=Dfjx7W1jKO8

Morozov, E. (2018b, March 31). After the Facebook scandal it's time to base the digital economy on public v private ownership of data. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2018/mar/31/big-data-lie-exposed-simply-blaming-facebook-wont-fix-reclaim-private-information

Mozur, P., Kessel, J. M., & Chan, M. (2019, April 24). Made in China, exported to the world: The surveillance state. *The New York Times*. Retrieved from https://www.nytimes.com/2019/04/24/technology/ecuador-surveillance-cameras-police-government.html

Mukherjee, M., Matam, R., Shu, L., Maglaras, L., Ferrag, M. A., Choudhury, N., & Kumar, V. (2017). Security and privacy in fog computing: Challenges. *IEEE Access*, *5*, 19293–19304. https://doi.org/10.1109/ACCESS.2017.2749422

Munn, L. (2018). Seeing with software. *Studies in Control Societies*, *2*(1). Retrieved from https://studiesincontrolsocieties.org/seeing-with-software

Narayanan, A., & Shmatikov, V. (2008). Robust de-anonymization of large sparse datasets. In Y. Guan (Ed.), *IEEE symposium on security and privacy* (pp. 111–125). New York, NY: IEEE.

Nikouei, S. Y., Chen, Y., Song, S., Xu, R., Choi, B.-Y., & Faughnan, T. R. (2018). Smart surveillance as an edge network service: From harr-cascade, SVM to a lightweight CNN. *Cornell University*. Retrieved from http://arxiv.org/abs/1805.00331

Ohm, P. (2010). Broken promises of privacy: Responding to the surprising failure of anonymization. *UCLA Law Review*, *57*, 1701–1777.

Pasquinelli, M. (2015). Italian operaismo and the information machine. *Theory, Culture & Society*, *32*(3), 49–68.

Roman, R., Lopez, J., & Mambo, M. (2018). Mobile edge computing, Fog et al.: A survey and analysis of security threats and challenges. *Future Generation Computer Systems*, *78*, 680–698.

Rouvroy, A. (2013). The end(s) of critique: Data behaviourism versus due process. In M. Hildebrandt & K. De Vries (Eds.), *Privacy, due process and the computational turn* (pp. 143–167). Abingdon: Routledge.

Sadowski, J. (2019). When data is capital: Datafication, accumulation, and extraction. *Big Data & Society*, *6*(1). https://doi.org/10.1177/2053951718820549

Shi, W., & Dustdar, S. (2016). The promise of edge computing. *Computer*, *49*(5), 78–81.

Shkabatur, J. (2018). *The global commons of data* (SSRN Scholarly Paper No. ID 3263466). Rochester, NY: Social Science Research Network.

Simon, D. (2018, November 27). Decentralized digital infrastructure: Towards digital commons. *Dezentrum*. Retrieved from https://www.dezentrum.ch/en/blog/decentralized-digital-infrastructure-towards-digital-commons

Simonelli, J. (2019, April 30). *DCD New York 2019 Q&A with Jim Simonelli, Schneider Electric* [Video file]. Retrieved from https://www.youtube.com/watch?v=m8iumNZycJU

Stack, T. (2018, February 5). Internet of Things (IoT) data continues to explode exponentially: Who is using that data and how? *CISCO*. Retrieved from https://blogs.cisco.com/datacenter/internet-of-things-iot-data-continues-to-explode-exponentially-who-is-using-that-data-and-how

Steyerl, H. (2016). A sea of data: Apophenia and pattern (mis-)recognition. *E-Flux*. Retrieved from https://www.e-flux.com/journal/72/60480/a-sea-of-data-apophenia-and-pattern-mis-recognition

Sverdlik, Y. (2018, July 31). When air no longer cuts it: Inside Google's AI-driven shift to liquid cooling. *Data Center Knowledge*. Retrieved from https://www.datacenterknowledge.com/google-alphabet/when-air-no-longer-cuts-it-inside-google-s-ai-driven-shift-liquid-cooling

Sweeney, L. (1997). Weaving technology and policy together to maintain confidentiality. *The Journal of Law, Medicine & Ethics*, *25*(2/3), 98–110.

Sweeney, L. (2001). *Computational disclosure control: A primer on data privacy protection* (Doctoral dissertation). Massachusetts Institute of Technology, Cambridge, MA.

Sweeney, L. (2002). K-anonymity: A model for protecting privacy. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems*, *10*(5), 557–570.

Taylor, L., Floridi, L., & Van der Sloot, B. (2016). *Group privacy: New challenges of data technologies* (Vol. 126). Cham: Springer.

Terranova, T. (2018). *Data mining the body of the socius*. Berlin: Staatliche Museen Zu Berlin. Retrieved from https://smart.smb.museum/export/downloadPM.php?id=5349

U.S. Congress. (2018). *CLOUD Act* (H.R. 4943). Washington, DC: U.S. Congress.

Valerio, P. (2018, October 31). To comply with GDPR, most data should remain at the edge. *IoT Times*. Retrieved from https://iot.eetimes.com/to-comply-with-gdpr-most-data-should-remain-at-the-edge

Wang, C., Zhang, Q., Liu, W., Liu, Y., & Miao, L. (2019). Facial feature discovery for ethnicity recognition. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *9*(1), 1–17.

Wang, H., Zhang, Z., & Taleb, T. (2018). Editorial: Special issue on security and privacy of IoT. *World Wide Web*, *21*(1), 1–6.

Warren, S. D., & Brandeis, L. D. (1890). Right to privacy. *Harvard Law Review*, *4*(5), 193–220.

Yi, S., Qin, Z., & Li, Q. (2015). Security and privacy issues of fog computing: A survey. In K. Xu & H. Zhu (Eds.), *International conference on wireless algorithms, systems, and applications* (pp. 685–695). New York, NY: Springer.

Zhang, H., Wang, Y., Du, X., & Guizani, M. (2018). Preserving location privacy in mobile edge computing. In X. You & C.-M. Chen (Eds.), *IEEE international conference on communications* (pp. 1–6). New York, NY: IEEE.

Zhang, J., Chen, B., Zhao, Y., Cheng, X., & Hu, F. (2018). Data security and privacy-preserving in edge computing paradigm: Survey and open issues. *IEEE Access*, *6*, 18209–18237.

Zhang, Q., Zhang, Q., Shi, W., & Zhong, H. (2018). Firework: Data Processing and Sharing for Hybrid Cloud-Edge Analytics. *IEEE Transactions on Parallel and Distributed Systems*, *29*(9), 2004–2017.

Zuboff, S. (2015). Big other: Surveillance capitalism and the prospects of an information civilization. *Journal of Information Technology*, *30*(1), 75–89.

Zuboff, S. (2019). *The age of surveillance capitalism: The fight for a human future at the new frontier of power*. New York, NY: PublicAffairs.

About the Author

Luke Munn uses both practice-based and theoretical approaches to explore digital cultures, investigating how technical environments shape the political and social capacities of the everyday. He is based in Aotearoa, New Zealand. His work ranges from data infrastructures in Asia to migrant surveillance in the Pacific and far-right cultures online. He has recently completed a doctorate at Western Sydney University on algorithmic power.