1

Introduction

Khurshid Ahmad
Trinity College, Dublin

Microblogging systems and social networking systems, sometimes referred to as types of social computing systems, generate enormous amounts of data that often contain specific information about persons, events and geographical locations. During exceptional circumstances, this data can, in principle, be invaluable – for example, in deploying scarce resources in times of emergency in order to save lives, rescuing and rehabilitating people and protecting property. Microblogs and social networking pages have become a mainstay of data dissemination in the public space and are used heavily by both citizens and authorities in times of crisis; they have been used, for example, by police forces for disseminating and seeking information about missing people, and by hospitals seeking blood. Twitter often carries the first messages of victims of atrocities (Paris 2015, Berlin 2016, London 2017 and Manchester 2017); Facebook is used to find missing relatives. Vast quantities of personal information floods microblogs and social networking pages. Although this data can help authorities to deploy resources where the need is greatest, it can also be used in wholesale surveillance. Social computing systems are now part and parcel of our interactions with others in good times and bad.

The ownership of the data in this space is disputed, and access to it remains a contentious issue. Social computing systems have all the hallmarks of a *disruptive* and *intrusive* technology, bringing unforeseen opportunities and challenges that cannot be faced by the existing ethical and legal order. The challenges related to the abuse of texts, images and spoken data available on social computing systems are all well-documented; the internet's ubiquity and transnational overreach can facilitate the appropriation of personal data and proprietary information across multiple social media platforms and national borders, anchored by multiple servers located in several countries, a scenario that inevitably raises jurisdictional issues in compliance and enforcement.

The scale of social computing means that its analysis – extracting relevant information from noisy data, checking facts, avoiding duplications, eradicating "fake news", recognising people and objects in grainy and shaky images – can only be adequately handled by yet another disruptive and intrusive technology, *artificial*

intelligence (AI). AI techniques can be more effective than experts in revealing connections between people and between objects. The ethical and legal issues raised by the use of AI have just begun; imagining the combined use of the two intrusive and disruptive technologies is a challenge the authors of this volume have researched severally and collectively over the past number of years.

Project Slándáil (2014–2017) was an EU FP7-funded project (project #607691) that leveraged social media data for emergency management. The technologies developed on Project Slándáil were supported by the legal and ethical research discussed in this book. As a follow-up to this project, Trinity College, Dublin has developed a plan to commercialise text and image analytic systems, gratefully supported by Enterprise Ireland (contract #CF-2017–0778-P) and the European Regional Development Fund (2014–2020).

1.1 A NOTE ON TERMINOLOGY

We all use, in one form or the other, systems that facilitate social interactions. These interactions include the establishment and sustenance of friendships or animosities, the exchange of goods and services at an individual or organisational level, the creation or destruction of communities (howsoever defined), the conduct or blockage of political debate, and individual or governmental crossing of societal norms. The term *social computing systems* is evolving, and social computing systems essentially comprise two types: social media systems and social networking systems (Parameswaran & Whinston, 2007; Wang, Carley, Zeng & Mao, 2007). According to the Oxford English Dictionary:

> ***Social media***. Noun. Websites and applications which enable users to create and share content or to participate in social networking. (In use since c. 2004.) [Exemplars of social media systems include Twitter, Flickr, Instagram, Snapchat, and YouTube]
>
> ***Social network***. Noun. A system of social interactions and relationships; a group of people who are socially connected to one another; (now also) a social networking website; the users of such a website collectively; cf. social networking n. (In use since c. 1845.) [Exemplars of social networking systems include the ubiquitous Facebook; more specialised systems include professional networking systems such as LinkedIn, personal networking systems for dating, and those for finding long-lost friends]

Social computing is used for technologies that enable people to store for free their (highly unstructured) data in expensive, large databases (usually meant for structured data). Social computing systems process the unstructured data by using advanced computer systems involving thousands of computers working together with state-of-the-art technologies such as AI and machine learning, and distribute the processed data or information to the world in microseconds. These technologies

tag unstructured data using details of the creator, the subject matter keywords, the machines used in generating the data, the geo-location of the creator and by annotating images and other seemingly innocuous metadata so that such data can be stored and retrieved systematically and efficiently (King, Li & Chan, 2009; Zeng, Wang & Carley, 2007). This involves the use of expensive and advanced information extraction systems but is, for users, all free.

Social computing systems "mine" the data of users, learning complex patterns within the data to mine more effectively. A symbiotic relationship exists between technology and data: larger datasets help systems to learn better, thus enabling better and even larger datasets to be mined.

Social computing systems connect disparate items of data in very complex ways:

- The **biographical data** of a person can be linked to their shopping habits, as displayed when they *crawl* shopping websites, or to inform them about private and public events.
- **Biometric data** is used by governments to plan and deliver health and social services, or to enforce immigration control.
- ***Merged* or *fused* biographical and biometric data** are used for rescuing and rehabilitating internally displaced people; disaster-struck people can be tracked by drones and by geo-locating their social computing messages in text and images.

Social computing systems connect disparate items of data together in ways sometimes referred to as *dataveillance*. There are two overlapping definitions here (Hu, 2015: 774):

(i) "In the intelligence context, it appears that 'collect-it-all' tools in a big data world can now potentially facilitate the construction, by the intelligence community, of other individuals' digital avatars. The digital avatar can be understood as a virtual representation of our digital selves and may serve as a potential proxy for an actual person".

(ii) "This construction may be enabled through processes such as the data fusion of biometric and biographic data, or the digital data fusion of the 24/7 surveillance of the body and the 360° surveillance of the biography".

The central question for us is why we, the citizen/users, make these systems so powerful? Two reasons come to mind. First, for small gains – free email, free software, free phone calls – and sometimes out of necessity – booking airline tickets, buying online, bidding in electronic auctions; we give away one of the most important assets we have, namely our biographical details. The second reason relates to the prolixity of legislation by many state agencies and the verbosity of service-provider forms, meaning that a vast quantity of biographical and biometric data is being continuously demanded of individuals. These vast silos continue to lose data, may contain data the silo should not have and can miss out on vital data.

Social computing systems are a classic example of disruptive technology and deft sales tactics. Disruptive technologies are paradigm-busting developments: a current technology (and some of its users) are discarded for a relatively untested new technology that creates its own, bigger user groupings. Disruptive technologies often emerge through a collaborative effort of like-minded, usually technocratic, persons and are marketed using nimble phrases such as "shareware" and "freemium". Most social computing enterprises have emerged during the past ten years or so.

Social computing systems have grown organically, and as such the ethics and legality of their use will continue to be debated by legislators and scholars. States and federations have struggled to adapt their legal systems to the developing technological systems; there appear to be few explicit rules, regulations, laws or directives for the development of these technologies, nor are there well-defined ethical frameworks for assigning rights and duties to the developers, end users and the individuals whose data is being acquired and disseminated.

EU Data Protection Regulations have just been released for consultation; Fair Use Legislation is lagging behind; national security organisations appear to have carte blanche for mounting surveillance on all known communications technologies while private security organisations are not far behind (Barnett, 2015; Bekkers, Edwards & de Kool, 2013; Galicki, 2015). Recent changes in EU law, such as the landmark developments in the EU's General Data Protection Regulation (2016/679) and the EU Law Enforcement Data Protection Directive (2016/680) together with the recently invalidated US-EU Safe Harbour framework and the EU-US and Swiss-US Privacy Shield Frameworks, approved in 2016 and 2017, respectively (Ni Loideain, 2016; Weiss & Archick, 2016), demonstrate the contested and evolving nature of the legal environment.

Social computing systems become even more powerful when coupled with *geographical information systems* and emerging *geospatial information systems*. In the OED, we find:

> ***geographic information system***. Noun. An information system which allows the user to analyse, display, and manipulate spatial data, such as from surveying and remote sensing, typically in the production of maps.

The general-purpose maps produced by cartographers comprise many themes – political boundaries at national and local levels, population distribution, vegetation and soils, transportation networks and elevation. The term "geospatial" is defined by the OED as follows:

> ***geospatial.*** Adjective. Of or relating to geographical distribution or location; esp. designating data associated with a particular geographical location; relating to or involving such data.

The introduction of geospatial data related to these themes, which is kept updated to minute levels – such as location-based details of individual houses, the number of trees on major road-networks, and census data at a very fine grain of detail – will provide opportunities and challenges that perhaps will be equal to or greater than those made available by web-search systems and their coupling with social computing systems (Groot & McLaughlin, 2000; Lo & Yeung, 2007). Social media systems comprise links to documents on the web, and news sources use social media and networking messages on their website. *Geospatial* search engines allow for the linking of data about place-based assets with other such data, and one can use social media systems to discover how to access and provide such informational data. The scope of the ethico-legal debate appears to be wider than simply respecting the factual and provenance of textual and image-based documents.

The United States is in the vanguard for exploring and exploiting the use of spatial data and has developed a web-accessible National Spatial Data Infrastructure (NSDI) to ensure "that place-based data from multiple sources (Federal, state, local, and tribal governments, academia, and the private sector) are available and easily integrated to enhance the understanding of our physical and cultural world".[1] The NSDI has been developed since the 1980s and is now being used as a backbone for crawling programs that will search and aggregate geospatial data in an unprecedented manner. Like the commonly available text (and some image) search engines, the advent of *geospatial search engines* (Bone, Ager, Bunzel & Tierney, 2016) will provide commercial, policy and public protection opportunities much as the earlier text/image search engines did, and will bring perhaps greater ethico-legal problems.

European countries have their own GIS and geospatial databases in different states of availability and development. However, more recently, the EU has created a policy framework and concomitant directives for the creation of an **i**nfrastructure of **sp**atial **i**nformation ac**r**oss **E**urope (INSPIRE) covering thirty-four data themes, ranging from the conventional themes found in most maps and GIS systems to data on sites of special scientific interest, human health and safety, atmospheric conditions, socio-economic data and so on (Bartha & Kocsis, 2011; Vandenbroucke, Zambon, Crompvoets & Dufourmont, 2008). The INSPIRE Directive was released in 2007, and the completion date of this large geospatial database for all twenty-seven EU countries is 2021 (at the time of this writing). The EU states have been mandated to make all this information available, at cost, to people across the EU and beyond (provided national security considerations allow for this).

The coupled use of GIS and geospatial data was initially planned for policy development covering areas vulnerable to flooding, fire, strong winds, poverty alleviation and economic and social development. With the advent of efficient, low-cost remote sensing equipment that generates data in real-time, such coupled systems have been used in disaster management, the monitoring of epidemics and logistics management, to name but a few applications. GIS and geospatial data manipulation have been extensively used by the military and security agencies to

monitor the movement of people, vehicles, arms and military machinery (Brewer & McNeese, 2003; Coleman & McLaughlin, 1998; Sui, 2008).

Social computing systems were initially used for crowdsourcing data for thematic maps, and human sensors were used to capture geospatial data (Heipke, 2010). Now, with the help of global positioning systems, we can track the activity of a person who is online and monitor mobile phones which have their "location finder" option on (Agarwal & Lau, 2010; Ryder, Longstaff, Reddy & Estrin, 2009). Tracking a person's communication equipment, mapping it onto geo-locational maps and observing its proximity to other people carrying mobile devices, facilitates monitoring of their movements and related activities. The computer systems of many law enforcement agencies are, as a result, filled with data resulting from legal or illegal surveillance of citizens going about their legitimate business (L. M. Austin, 2014; Greenwald, 2013; Kerr, 2002; Lyon, 2006; Pell & Soghoian, 2014).

1.2 SECURITY, PRIVACY AND DIGNITY DURING AN EMERGENCY

Systems are being developed that can process texts and images in conjunction with geospatial data in order to extract information that can be used during emergencies, using social media as a bi-directional communication aid between authorities and vulnerable citizens (Alexander, 2014; Simon, 1982). The authors of this book were involved in the development of such a system (Ahmad, 2017). In tandem with the project's technical work, we conducted an ethical analysis to ensure that the inviolable rights of citizens related to privacy and dignity, the concomitant proprietary rights of the individual to his or her social media texts and images, and the constitutionally mandated duty of the authorities to protect lives and property, were all given due consideration.

One important design consideration for us, therefore, was to ensure that the social-media informed emergency management system was licensed according to the current laws that allow safe use of the internet, protection of data available on social computing systems and, above all, that the rights of the citizens, whose life and safety are in peril during a disaster, were not violated whilst the emergency managers performed their mandated duties. We have developed a *licence* for the use of such systems wherein the user agrees to abide by the law of the land and international laws for an authorised use of the system.

A number of questions guided our investigations. Is there an ethical basis for an institution to examine an individual's interaction with others that is within the bounds of ethical norms and is permitted by the law? Can enterprises and agencies provide a legal basis for surveillance of citizens in cases where monitoring and harvesting of data has the potential to contribute to the public good? Both in the provision of opportunities and the posing of challenges, one question keeps recurring: who "owns" the data, those who generate it or those who store it and make it available on demand? *Social Computing and the Law* explores these questions,

presenting a comparative transnational overview of the legal ramifications of harvesting social media data on the internet in the United Kingdom, the European Union and the United States.

Our team of lawyers worked with four emergency managers (two police forces, one in the UK and the other in Ireland), two civil protection organisations (in Italy and Germany), and two software companies (based in Germany and Italy), together with experts in human rights and ethics (in Ireland and Italy) and computer scientists (in Irish, German and UK universities) to chart the ethical and legal landscape surrounding the use of social computing data in exceptional circumstances. Our three lawyers – one a professor in internet law, another in copyright and data protection law, and the third in human rights laws and conventions – researched these areas for the use of social computing in exceptional circumstances; they were counselled by a professional law firm in Ireland and by academic experts in value pluralism.

We found that disaster management represents an invaluable case study for processing social computing data in exceptional circumstances, demonstrating the way in which information ordinarily considered private may be monitored and harvested for limited time periods in the interests of the public good. Social media allows an unprecedented look into the activities and environments of individuals and communities, presenting information that can be invaluable to emergency managers in each phase of disaster response. The potential for abuse of this data, however, is considerable, and it should and can be managed in the first instance by having a licence governing terms and conditions attached with the sale and use of a social media monitoring system that can be used in an emergency.

Social Computing and the Law presents key findings of our research on internet governance, data protection, copyright and human rights, together with a discussion of key motivating ethical issues concerning the problems of competing rights and the delegation of power to authorities. We look at ethical deliberations on how to reconcile the conflicting needs of two groups and consider the governance of the internet, as exemplified in relevant legislation and protocols, to ascertain the rights and duties of the vendors of these systems. We explore how copyright protection may be of help in protecting the rights of the citizens who generate data. We consider how international human rights laws and protocols provide an overarching umbrella for the protection of privacy and dignity as cardinal principles of governance. The incorporation of international human rights law within national legislative machinery can counterbalance the tendency of nation-states to curtail citizens' freedom.

The synthesis of these three legal analyses has led us to develop a template software licence drafted by professional lawyers (see Appendix A), with an accompanying legal checklist (see Appendix B), that provides a legal framework for the safe and transparent deployment of social computing systems, especially during exceptional circumstances. The book concludes by presenting these documents, showing

how a practical and enforceable legal agreement can incorporate the ethical concerns and competing rights surrounding the use of social computing data in exceptional circumstances.

1.3 OUR CONTRIBUTION: DISASTERS, TECHNOLOGY, LAW AND ETHICS

Our book sets itself at the crossroads of several rapidly developing areas of research in legal and global studies relating to social computing. The advent of the internet has influenced all manner of legal and regulatory frameworks, including the protection of individual rights such as privacy, data protection and intellectual property; the explosion of social media has further expanded the challenges for such fundamental rights. *Social Computing and the Law* highlights complex legal challenges to these rights and considers how public emergency responders could legally appropriate content on social media platforms for emergency and disaster management. In the process, the book makes significant contributions to the ongoing debate on the corporatisation and commodification of user-generated contents on social media and the extent to which these can be legally and ethically harnessed for public emergency and disaster management.

Recent work on the societal impact of social media in exceptional circumstances has led to position papers on the ethical and legal aspect of using social media; see, for example, Hiller & Russell (2017). There is a growing discussion surrounding the ethical and legal challenges of using data gathered during a disaster as well as on the effect of EU Data Protection Directives, the 2016 version thereof (Rizza, Büscher & Watson, 2017). Our book contributes to this evolving literature by providing an extended exploration of the relevance of internet law, copyright and data protection law (with a focus on the changing legal landscape of the EU) and human rights laws to the propriety of using personal data in crisis management. Our checklist of legal obligations and ethical issues (and the accompanying discussion) will provide a valuable contribution in this regard.

Existing monographs on the use of social media in emergencies show a distinctive focus on emergencies affecting states in connection with armed conflicts, terrorist threats and similar political predicaments (see, for example, Gupta & Brooks, 2013; Nissen, 2015). Natural disaster scenarios have remained comparatively less studied, despite the pervasiveness and societal impact of such situations, and the significant role that technology played in them. The reflection on the appropriate use of big-data analysis and data mining in connection with natural disaster scenarios has just begun. In addition to the enormous positive potential for emergency response managers arising from such a shift in natural disaster management, legal and ethical risks may arise if disaster response operators fail to consider privacy and data protection as crucial human rights concerns.

The book benefits from the multifaceted work produced in the Project Slándáil, which combined theoretical and practical approaches; as an interdisciplinary work

that draws on ethics, law, computer science and disaster management, it has very few competitors. Moreover, major works in the field of ethics of disaster management make few if any references to the ethical and legal challenges of social media harvesting – for example, in the area of disaster relief (Caron, Kelly & Telesetsky, 2014) or in disaster response (De Guttry, Gestri & Venturini, 2012). Works on ethics and social media explore key questions of *data/information* and *ethics* in a broad, general setting (Ess, 2013; Floridi, 2013; Taddeo & Floridi, 2017; Vallor, 2016) rather than deal with the special case of emergencies – something we have attempted to do (Jackson & Hayes, 2016).

The existing legal literature takes a rather generalised approach to the treatment of the preceding issues (De Franceschi & Lehmann, 2015; Lloyd, 2017; Stewart, 2017) and seldom has examined in detail the possible legal challenges posed to the altruistic uses of commodified user-generated contents on social media by public emergency responders. Works on data protection, copyright law and privacy law are many, but these either tend to focus on one aspect of the law or fail to attend to the situations of crisis or emergency management [Bently & Sherman, 2014; Citron, 2014; Mayer-Schönberger & Cukier, 2013; Reed, 2012; also D. K. Citron's *Hate Crimes in Cyberspace* (2014)].

The work carried out in Project Slándáil in building an emergency management system that receives actionable data (namely, data that is filtered, relevant and geo-located) from social computing systems involved extensive use of major branches of AI. Our system accepts data from social computing systems and from other digital streams comprising newspapers. The data is processed for information sharing with due regard to its copyright and due regard to the processed information for saving lives and property, but it is also capable of revealing identities of people and places. The processed information is analysed further to its context: information about the location (of a disaster) and the situation in which people and places may be. Finally, the processed information is aggregated for issuing warning, and for this knowledge for mitigating the impact of a disaster. In the final phases the system ensures that it is not distributing copyrighted material and that it is protecting the identities (see Figure 1; details are in Zhang, Kelly & Ahmad, 2016). These included natural language processing and computer vision systems, together with advanced geographical information systems. The use of these systems can generate information that can potentially violate the privacy and dignity of individuals, and concerns about the ethics of using AI systems has recently been aired in, for example, Emanuelle Burton et al.'s "Ethical Considerations in Artificial Intelligence Courses" (2017). Burton and colleagues make a concrete contribution to this literature by dealing with the real-world integration of AI technology into disaster management.

Engagement with legal practitioners adds a practical dimension to our work, which endeavours to translate ethical and legal principles into enforceable norms. The book, at the intersection between human rights law, internet law and disaster law, provides a holistic, thoughtful and timely contribution to a fast-growing debate.

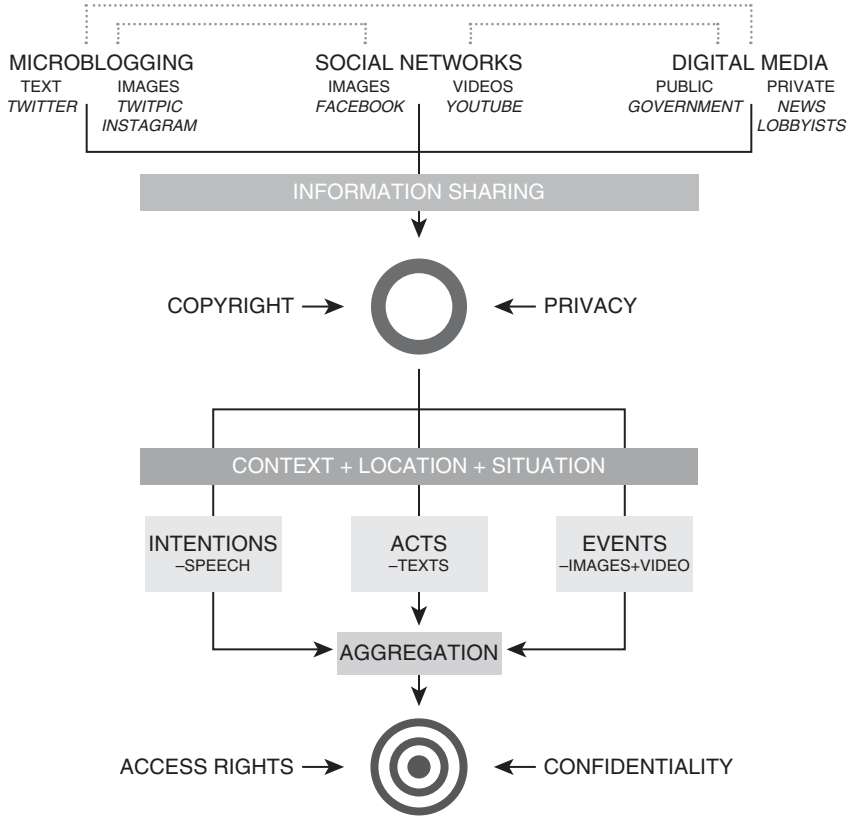

FIGURE 1 The complexity of using social media for emergency management, including sources and types of data. Acknowledgement of the legal landscape, including copyright and privacy, affect both the collection of data and the sharing of actionable information to emergency management teams. (Image courtesy of Project Slándáil – EU-FP7 Project #607691)

1.4 STRUCTURE OF THE BOOK

Chapter 2

We focus on social computing systems that deal with texts and images and will refer to geospatial information in passing. We begin with a discussion of the nature of the internet – the interconnected system of networked computing assets including processors, databases and graphical user interfaces – and discuss laws and conventions governing its use, particularly in the EU. The internet's ubiquity and transnational overreach could facilitate the appropriation of personal data and proprietary information across multiple social media platforms and national borders, anchored by multiple servers located in several countries. This inevitably raises jurisdictional