4. Mobilising open data

Introduction

This chapter addresses the dynamics of the ways in which open data is being mobilised in society. The previous chapters noted how the position of scientific knowledge in wider social and economic life has changed in late modernity, and explained that data and open data are being discussed by governments and civil society organisations in ways that focus on the possible social benefits of open data. These two areas – the changing role of knowledge in society and the possible benefits of open data – should be viewed in relation to each other, because the aspirations for open data are often couched in terms of the role that knowledge plays in society. Furthermore, the possibility of open data, combined with changing senses of the position of knowledge, are utilised within various visions of a future, more knowledgeable society. Within these discussions, there is recognition of the potential benefits to society as well as an acknowledgement of what needs to be done to ensure that open data is produced and used in responsible ways. An overarching theme in these discourses is that of 'open' – both in terms of open data and open society.

To address the dynamics of mobilising open data, the chapter first outlines the various main ideas about what constitutes open, and the value of open as understood by open data advocates. Second, it identifies the ways in which open data advocates are working in relation to the characteristics of social movements. Given that the value of open is key idea or concept within notions of open data and open society, the next section considers of the value of open in knowledge production. The chapter then outlines the history of openness as part of WWW culture, before looking at the configuration of actors that are combining to create a movement for open data. Next, the chapter addresses the ways in which ordinary people can engage with data, and points out that the development of interpretive communities is important in making data useful for wider populations.

Summary of the overarching context of a movement pushing for open data

Advocates of open data – each having visions and values of open data – are organising themselves with the aim of mobilising open data in society. This

mobilisation is based on their respective values and visions, which are not necessarily based on an explicit notion of a knowledge society, although they each work with ideas about a knowledgeable society. There is a wide range of understanding about what is meant by open, and how it might best be mobilised in data, in science, in knowledge production, and in society. Questions about what open means in science and in society have been debated historically and are re-emerging in relation to open data and knowledge society. The debates and visions many well inform change; however, as Stehr (1994) notes, the transition to a knowledge society is not being achieved in any planned or coherent manner; rather, the realisation of a knowledge society is occurring through an array of actions in relation to various uncoordinated institutional frameworks. They are linked through various types of social mobilisations that are clustering around the theme of open in relation to data, and all relate to ideas about open science and open society. As the previous chapter explained, government policymakers and civil society organisations are seeking to develop open data. Although the policy push is a key feature in mobilising open data, civil society organisations are a strong mobilising force. They play a key role by acting as a data movement and, to some extent, a social movement. Social movements are instrumental in shaping the discourses around causes, in bringing people together to push for change, and in connecting policy sensibilities with public sensibilities. This is important because, in order for open data to produce the transformation into a knowledge society, citizens using data in their daily lives must be involved, just as much as knowledge producers and government services. This discussion links explicitly with RECODE because of the ways in which scientists and other stakeholders, who are often outside the open data movement, must be mobilised in order to achieve the policy and civil society goals of a knowledge society.

Understanding the mobilisation of open data as a movement

In general terms, the concept of a social movement is rooted in theories that seek to address various forms of collective action.[1] Although there is debate about different types of social movement, the 'new social movement

1 There are several approaches within the study of social movements, such as political process theory, and resource mobilisation theory.

theory'[2] – in the context of Castells' (2001) digital networked society – provides a lens through which we can understand the ways that social movements are operating in the context of open data. A distinctive feature of contemporary social movements is that they involve a 'lifeworld' focus, where debate and communication serve to create a normative consensus (Habermas 1984, 1987). In part, this is a reaction to the fact that economic and political institutions are increasingly interfering in lifeworlds. These intrusions generate responses that are organised though social movements based around issues such as quality of life, democratic participation and identity (Staggenborg 2011). Although the focus on culture is often seen as distinct from structure, Polletta (2004) argues that these need not mutually exclusive. Rather, when culture is understood as 'the symbolic dimensions of all structures, institutions and practices' (*Ibid*., p. 100), then social movements are engaging with both structural and cultural issues. This focus means that social movements in contemporary society can be defined as 'purposive collective actions aiming at the transformations of values and institutions of society' (Castells 2001, p. 138). Open data spans across both cultural and structural areas of contemporary society because knowledge both structures social life and is culturally shaped by social relations.

Another feature of contemporary social movements is that they use the flexibility of the WWW to organise action and often manifest themselves via digital platforms and networks (Castells, 2001). Given this context, Melucci (1996) argues that social movements are no longer collections of relatively stable organisations, or even unified actors, but instead are often fluid networks that can foster collective action as and when needed. Social movements can then be characterised by their fluidity and flexibility as well as their focus on particular issues or values that might be distributed across networks. Melucci (1988) writes that, 'to understand the way social movements are constructed requires looking at the formation and maintenance of the cognitive frameworks and social relationships that form the basis of collective action' (*Ibid*., p. 331).

The dynamics of these movements means that relationships are often formed within submerged networks in which new collective identities

2 There is some debate that questions whether there really is a 'new' social movement theory, because some of the issues, such as labour, predate post-industrial and information society issues (Cohen and Rai 2000). Nonetheless, commentators such as Klandermans (1986) argue that new movements are emerging from a range of issues and grievances rooted in post-industrial society and in information society developments. Examples of these include the peace movement, the environmental movement, and the women's movement.

or identifications with an issue are formed. From this type of activity, activists generate new cultural models and symbolic challenges (Mueller 1994, pp. 247–248). Activists' perceptions of, belief in and emotions towards the cause, along with their adherence to a set of values, all play into the way that social movements emerge and act (Staggenborg 2011). Developing and framing an issue or issues under question involves interaction and mobilisation, which occurs through particular episodes of contention. This leads McAdam *et al.* (2001) to assert that attention needs to be paid to how actors attribute threats and opportunities, how they appropriate mobilising structures, construct frames and meanings, and innovate collective action tactics.

An analysis of the way that civil society organisations frame open data and build networks reveals a striking similarity to the way that social movements work (see Chapter Three). For example, activists lobbying for open data are seeking to mobilise large-scale changes and working across everyday networks of personal contacts as well as organisational structures of data producers and providers. There is also some indication of a collective identity emerging across the various networks that are pushing for open data. Currently, this identity is not highly formalised but is, nonetheless, recognised as people who are 'pro' open data, whether in research, government or civil society contexts. Some aspects of the networks are to some degree submerged, such as the work of repositories and data archives. There is also a drive to develop new ways to support open data, such as ecosystems, new institutional guidelines and new data practices. These types of activities are advocating a whole new set of values around data: in particular, proposing that data should no longer be owned by any one institution or person, but be publicly-owned instead. Furthermore, some new constituents in the ecosystem are emerging, in terms of institutional relations, new data curation services and new types of expertise and practice, such as data management plans, which have to be managed and implemented by those who may be outside the recognised social movement. These developments are part of the way that civil society organisations are pressing for open data, but they will only be accomplished if those other stakeholders are successfully motivated and mobilised. This mobilisation is also built on an ideology of openness that is a distinctive feature of early digital culture. Even if this has been partially undermined through the commercialisation of the WWW, openness still remains a defining feature of digital culture.

Openness as a value: Society, science and the World Wide Web (WWW)

When discussing open data, it is important to understand what is meant by open – not only in relation to data, but also in relation to knowledge production and society. There are many uses of open and openness, which are often related to notions of freedom and what is free. Open can refer to open society as well as the way that open is practised in a range of areas, such as free speech and free software. This notion of open is also used in expressions about freedom, such as freedom of expression, freedom of the press and freedom of movement. These are expressions of aspects of open society and the ways in which open society is negotiated. The notion of open is important in assessing the way that open data and open science or knowledge production are related in transforming society into a knowledge society which, in its idealistic sense, creates an open society.

Debates about open society were being held as long ago as the world of classical Athens where, for example, Pericles spoke out in favour of an open society. Pericles called for a society in which citizens were equal before the law and had influence in society (Brin 1998). However, this early vision was soon crushed after the Peloponnesian War (431–403 BC), and Plato later questioned the wisdom of having an open society and democracy (Popper 1966). The influence of Plato and his followers meant that the idea of an open and free society was not being advocated until the philosophical work of Locke during the Enlightenment, which was then developed in America by politicians Jefferson and Madison. These early notions of an open society were not, however, inclusive societies, as they incorporated slavery and colonialism, which had political structures that generated oppression and imprisonment for some against the rights and freedoms of others. Reflecting on Nazism and Stalinism, Popper (1966) explores the way that Plato's hatred of empiricism and democracy flowed through into Hegel's ideals, then onto Nazi officers and Marxist-Leninist commissars. Attempts to overthrow these oppressive regimes, from the 1940s to the fall of the Berlin Wall in 1989, resulted in a period where the idea and reality of freedom seemed to hang in the balance.

Popper (1966) highlights that notion that open society is fragile. Writers such as Orwell and Huxley addressed the ways that new technologies, surveillance and management programmes threatened open society. The struggle for open society continues today, and is on a global scale – in both old and new ways. Thus, slavery still exists in some forms, whereas digital technology closely monitors individuals within a surveillance society. The

debate about what open society might mean in a digital context is once more raising concerns about freedom and democracy. In this context, freedom is debated in terms of who has what knowledge and the belief that there needs to be a transparency about what knowledge actors have of each other. In the popular writings of David Brin (1999), this is termed as working towards a 'transparent society'.

This brief overview provides some evidence that the notion of open, and its related concept of freedom, has a long history with a broad application. In science, open is mostly used in debates about open science within the institution, as well as being part of wider society. Chapter Two showed how science is embedded in social relations, negotiating what open means for its work and considering how it relates to society more broadly. If we take science to mean the systematic pursuit of knowledge, then it plays a role in making the world knowable. As Chapter Two outlined, the position and role of science is changing, in line with changes in society, and is doing so both philosophically and empirically. Fuller (1999) notes that science as the pursuit of knowledge underwent a range of changes during the twentieth century. However, even including shifts such as Mode 2 knowledge production, Fuller stresses that the political rhetoric which positions science in society remains largely unchanged. He writes that science is still governed by a self-selecting group who decide who can qualify as scientists through examinations and by determining what is deemed to count as knowledge, through publication strategies. The position of science and its authority as a self-regulatory institution is a generally-accepted normative one. Even when cases of research fraud are discovered by the research community, these are seen to validate the critical stance in science. Science's authority means that it is listened to by policymakers, industry and ordinary people who not may fully understand the science, which, in reverse, means that scientists may not fully understand the context in which science is applied. This 'mutually tolerable ignorance' is something that is established and accepted (Fuller 1999).

A knowledge society is partly based on this normative position of science, with an assumption in the normative ideal that science is an open community. Although imperfectly realised, the ideal of the open society of science remains dominant in policy areas. This links with Steve Fuller's argument for what he terms a 'welfare economics of science' or 'knowledge policy' that will ensure equal and informed access to data and knowledge (Fuller 1988, 1993, 1997). Fuller's argument for an open science is shaped by republican values, as espoused by Popper's stance on open society. A republic of science or open science seeks to position itself between the

excesses of communitarian and liberal approaches within a society-science relationship. For example, societies with a strong communitarian ethos may censor research that could be used in negative ways in policy, such as the possibility of creating a 'racial science that could be used to develop a political movement like Nazism' (Fuller 1999, p. 12). In terms of liberal societies, there is a threat to the integrity of science in the way that market values intervene in free inquiry. In this context, free market and free inquiry are seen as the same thing, so any research can be undertaken as long as there is sufficient money to do it. Given that the cost of research is high across the range of disciplines, there is a tendency for research that attracts funding to develop, which may result in some areas of study being under-investigated. This may lead some researchers to seek funding from private investors for their work. There is also a lack of questioning the value of some well-funded research programmes, such as high-energy physics. These are the types of issues arising from the way in which 'science functions in society and this impacts on what kind of knowledge is produced, as well as how such knowledge may be used' (*Ibid*., p. 13).

To ensure that science can act as open science and as a republic of science, some basic conditions must be met which relate to the practice of science itself and to the ways that science and scientific knowledge acts in broader society. The conditions that underpin open science are that:

– People's opinions might change for the better as a result of hearing opposing opinions.
– People need not fear the consequences of their expressed opinions on their material well-being.
– There is a 'public good' or 'civic ideal' to which people may appeal in deliberation, which transcends specific individual and group interests (*Ibid*., p. 15).

If republican science policies is the ideal, then those policies will seek to make sure that everyone is materially secure enough feel confident in expressing their own opinions. Here, Fuller notes (drawing on Popper) that, if someone can express their thoughts with impunity, then his or her ideas can be judged by others in an open way (*Ibid*.). This is significant because it highlights the importance of being able to speak out. This is not always an easy thing to do, and research has shown that, even in open contexts, many people still feel deferential to a hierarchical order or they may fear humiliation (Elster 1993), both of which will deter them from stating their opinions. Therefore, both material and psychological conditions have to be met so that science can function in an open way.

Another factor in making science open is the strength of a civic ideal and/or public good and the ways that they are understood. To be able to address different interests in society, there needs to be some overarching body or set of principles that can assess the impact of one set of interests on wider society (Fuller 1999). The notion of special interests assumes that there is some established civic ideal against which different arguments and developments can be debated in terms of a public good. In this way, special interests within a larger collective are valued – not in their own right – but only valued by the extent they can contribute to the wider population. This is different to intellectual property, which seeks to protect information owners being forced to make their data freely available. Fuller (*Ibid.*) argues that this is how the epistemological distinction between pure and applied knowledge is transformed into an economic choice between public good versus intellectual property. There is also an internal assumption of a commons within science, which is seen in the way in which scholars do not question scientific theory by using reasoning and data from another field, discipline or sub-discipline outside their own. If a scholar does want to challenge these norms as well as any scientific principles, then they need to propose their new ideas openly, so that they can be publicly scrutinised. Fuller (*Ibid.*) argues that the concept of a civil ideal and public good generates an external boundary to the political and scientific endeavour. This enables internal changes to be followed, noted and questioned, and these types of processes create scientific conventions (Popper) or paradigms (Kuhn).

These processes are in play within the social relations of science that have political and economic influences internally on science as well as externally on wider society. This is significant because it affects what kind of knowledge is produced, the way that knowledge is produced and how knowledge is shared within wider society. Given that the normative idea of science still has authority and that knowledge society is based on ideas about the role of data in society, then the way that science functions affects how knowledge is produced and shared. Theorists of science such as Fuller (1999) assert that there are different models of science – some more open than others – which shape the possibilities of realising open science. Underpinning this, Fuller argues, is the principle of 'the right to be wrong' (*Ibid.*, p. 4). Therefore, in considering open data and the various open data movements and perspectives, it is necessary to ascertain if the material conditions are there to support openness. This means ensuring that scholars, scientists, policymakers and citizens have the right to be wrong and that there are sufficient checks and balances in both the production of

knowledge and the use of knowledge. As this discussion shows, openness in science takes a particular form based on the mores of scientific communities and the ways that science is practiced. Furthermore, the ways in which science interacts with society also shapes the way that research data can be made publicly available. Thus, a cultural change in the practice of science will be required to make scientific data open (see Chapter Six for more detail on this).

The development of open data and open science is partly being enabled by digital technology, with some ideas about open and notions of freedom being inherent. Berners-Lee, the inventor of the WWW, brought together hypertext and the internet to build the WWW, and CERN released the first browser over the internet in August 1991. Berners-Lee argues that, throughout the Web's history, there have been parallels between technical design and social principles (Berners-Lee 1999). He designed the Web on universalistic (with lower case u) principles, to build an environment that would enable people to think and discuss diverse issues from a range of perspectives in an open and accepting way. This informed the development of decentralised systems of computers, knowledge, and people. Berners-Lee's values provide a narrative that focuses on the forms of participation in the WWW, in which:

> hope in life comes from the interconnections among all the people in the world. We believe that if we work for what we think individually is good then we as a whole will achieve more power, more understanding, more harmony as we continue the journey. We don't find the individual being subjugated by the whole. We don't find the needs of the whole being subjugated by the increasing power of the individual. But we might see more understanding in the struggles between these extremes (*Ibid*., p. 228).

Berners-Lee (1999) understands freedom of the internet in two ways. First, freedom is experienced in terms of sending any content anywhere in the network in packets. Second, it provides a freedom of association which is based on mutual respect with an ethos of collective endeavour that goes beyond singular individual effort to build for the common good in ways that are unconstrained by bureaucratic regimes. This vision informed virtual communitarians who sought to use it to generate egalitarian and alternative communities. Their culture generated a context in which the internet moved beyond its specialist employment to more general social use. Thus, those early users of networked computing outside of university or hacker

environments created virtual communities, using the term popularised by Rheingold (1993).

Castells (2001) notes that these communitarians contributed to the shape and evolution of the internet, including its commercial manifestations in decisive ways – for example, the earliest Bulletin Board Services (BBS) in the San Francisco Bay area and the work of the Institute for Global Communication (IGC) – focused on socially-responsible agendas such as protecting the environment and preserving world peace. IGC established the first women's computer network (La Neta), which was used by the Mexican Zapatistas to build international solidarity on behalf of 'Indian Communities'. Other community networks, like Schuler's Seattle Community Network or the Digital City Amsterdam, sought to renew or enhance citizen participation. Another historically specific use of internet-based networks was the way that Russian academics used it to organise activities for democracy and freedom during the perestroika period of dismantling the Soviet Union.

The history of the internet and WWW is tightly related to a range of open movements in social terms, which are based on the WWW's open architecture – including open source, open hardware and open content. The open data movement forms part of this overall open movement ethos. The distinctive feature of the open data movement is that it focuses on data, asserting that data should be freely available for reuse and republishing. As noted in Chapter Three, the overall drive and movement itself is made up of a variety of groups with a range of data-related backgrounds, which are self-organising and aim to promote and facilitate open data. A range of civil society movements are involved, such as the Open Data Foundation (ODF) and the Open Government Group (OGG). This push for open data is fully supported by Tim Berners-Lee, who is calling for 'raw data now'. He explains that he has moved from an initial focus on openly sharing documents to believing in the need to share raw data. Berners-Lee (2012) says that the open data movement's position focuses on how open data can interconnect and join data to summarise and compare, to monitor, extrapolate and infer.

Open in this context focuses on how open data interplays with knowledge generation and how this can be facilitated by the promotion of a robust commons that will enable anyone to participate. There are practical aspects to realising this vision, which include maximising interoperability to cover a wide range of data, systems and licences. The ODF expresses this by advocating that 'knowledge is open if anyone is free to access, use, modify, and share it – subject, at most, to measures that preserve provenance and openness.' As noted previously, the ODF argues that this core meaning of open data matches the open source definition used in software, and is

synonymous with 'free' or 'libre' as defined by Free Cultural Works (https://creativecommons.org/freeworks/). The ODF's definition of open data was initially derived from the open source definition, which in turn was derived from the Debian Free Software Guidelines. Given this background, the term 'open data' as defined by ODF is used to denote the item or piece of knowledge being transferred. There is a level of regulation about this use, which is mainly operated through licences, as discussed in Chapter Three.

The configuration of an open data movement: The characteristics of social movements and actors in mobilising open data

As Chapter Three outlined, a range of organisations are actively promoting open data. There is also a strong push towards open data from governments and global regional actors such as the G8's Open Data Charter (https://www.gov.uk/government/publications/open-data-charter), the United Kingdom's open government data agenda (data.gov.uk), the United States' open data agenda (www.data.gov), the EU's open access to research data policy (EC 2016) and the OECD's Open Government Data project (OECD 2016). The G8, for instance, has high hopes for how data can be used in society and how open data could mobilise a knowledge society. It sees open data as an untapped resource with huge potential to encourage the building of stronger, more interconnected societies that better meet the needs of their citizens and allow innovation and prosperity to flourish. To foster the development of open data, the G8 has agreed to follow a set of principles that will provide a foundation for access to, and the release and reuse of data made available by G8 governments (see Chapter Three). The UK and US open government data initiatives are similar in that they focus on the ways in which open data can benefit citizens generally as well as finding new ways to facilitate innovation in both commercial and non-commercial settings. The G8 (representing many governments) recognises how important diversity is for stimulating creativity and innovation, believing that the more people and organisations use data, the greater social and economic benefits will be generated – again, for both commercial and non-commercial uses.

The basis of this type of action is a vision of open data that is largely based on a range of civil society organisations. The open data movement is rooted in the culture of freedom and openness that is at the heart of WWW culture. This continues today with Tim Berners-Lee still campaigning and now pushing for open raw data. This momentum and drive for change is also enacted by the civil society organisations which are seeking to mobilise

open data. These organisations often comprise elements of social movements in that they are framing the debate and agenda, they are garnering support from a range of actors, and they are articulating the opening up of data in both cultural and structural terms. They are addressing the cultural issues around opening up data by focusing on perceptions of how data should be shared and on the practices for sharing data. They are also addressing structural issues by pushing for institutional change to support open data as well as detailed changes at the institutional level, for instance, the legal, regulatory and ethical aspects of data. Each of these are relevant for all types of open data, including open research data, and form the 'grand challenges' that were examined within the RECODE project and presented in more detail in Chapters Five to Nine.

Although there is a common focus, the movement for open data is made up of a number of advocates based in civil society. There are different groups with varying foci and perspectives. As well as the key actors mentioned in Chapter Three, these include the Open Data Institute, the World Wide Web Foundation, the World Wide Web Consortium (W3C), the 'Web We Want', and the Open Data Research Network. All of these organisations are addressing open data and, although each considers open data in general terms, each organisation also considers how open data can be used in various aspects of social life. The 'Web Index', an organisation within the World Wide Web Foundation, focuses on measuring the Web's contribution to development and human rights at the global level.

Part of the way that open data advocates are working and joining together to make an open data movement is lodged in their respective focus on how to make data open, each arguing that a range of technical, institutional, legal and social factors need to be addressed in order to operationalise open data. This aim is based on a range of sub-themes, which inform and underpin the requirement of what is termed 'open works' or open data. To support the responsible sharing of data and making data open, one actor, the ODF, created sub-themes seeking to facilitate the distribution of open works. The sub-themes include open licensing, as ODF, in line with other open data advocates, asserts that open data must be available under an open licence and that any additional conditions accompanying the work (such as terms of use, or patents held by the licensor) must not contradict the terms of the licence (see Chapter Eight for a full discussion on this).

As Chapter Three discussed, open licences include a range of conditions and permissions that are based on key sub-themes found in the open data movement. These include free use and redistribution of the licenced work, including sale – whether on its own or as part of a collection made from

works from different sources. The theme of modification is also evident in licences, as are the creation of derivatives and the distribution of such derivatives. Another sub-theme is the demand that data should be available to be freely used, distributed, or modified separately from any other part of the work or from any collection of works in which it was originally distributed. The sub-theme of compilation also focuses on distribution, pointing out that open data can be used alongside other distinct works without placing restrictions on these other works. One major sub-theme is the belief that open data should be based on a non-discriminatory agenda, in that licences and general accessibility must not discriminate against any person or group. Attribution is another important sub-theme, insisting that open data users should give credit to the contributors, rights holders, sponsors and creators.

A further sub-theme addresses the issue of access, which states that open works and open data should be available as a whole and at a reasonable one-time reproduction cost, preferably downloadable via the internet without charge. Part of the accessibility sub-theme is that open data should be provided in open format, which means that the data must be provided in a convenient and modifiable form to ensure that there are no unnecessary technological obstacles to people trying to exercise the licensed rights. Specifically, data should be machine-readable, available in bulk, and provided in an open format (i.e. a format with a freely-available published specification which ensures there is no restriction – monetary or otherwise – on its use) or, at the very least, can be processed using just one free/libre/open-source software tool.

The way in which the various actors are operating in the push to mobilise open data is based on their definitions of open data. Chapter Three discussed various definitions of open data, the principles of those definitions, and the ways that open data can be used in practical terms. Social movements around open data operate in a similar way to that of others, in that several organisations each have a distinctive focus, but act together under the umbrella of open data to mobilise open data across many social areas. Thus, for example, OGD focuses on government data, ODF looks across a range of data but focuses on developing training for citizens, and the EU is pushing the commercial sector to open up big data.

These policy and civil society actors are working in the same mobilising space so, in that sense, are forming a social movement. The combination of these actors is acting as a network among particular domains of data, social constituencies and contexts to mobilise open data in wider society. The key principle of open data unites the distinctive actors, with each acting

in relation to their respective social constituencies to bring about change. This collaborative action is generating the conditions for data to contribute to a transformation to a knowledge society.

Open data in wider society: Citizens and organisations access and use of data

Although open data movements are seeking to develop open data, there is still the question of how ordinary people and organisations (public, private and third sector) can access and engage with open data. In order to mobilise change towards knowledge society, data literacy needs to improve across the population. It also requires an awareness and better understanding amongst people and organisations of how to unlock the value of open data. This involves education about data literacy and ways to interpret and use data as well as fostering good practice by providing data in machine-readable formats to empower a future generation of data innovators. The commercial sector is focusing its attention on how to realise the value of open data. In the US, for instance, the Chamber of Commerce Foundation is working with New York University's Governance Lab to explore how the value of open data can be realised in the commercial sector (see Chapter Nine). The attention in this context is on data-driven innovation, seeking to understand how open data and big data can be used for innovation (US Chamber of Commerce Foundation 2014).

In board terms, the US Chamber of Commerce argues that four kinds of open data drive innovation: scientific, social, personal, and governmental. There is an assumption that researchers work collaboratively with scientific data and that this is driving forward knowledge in scientific terms. In the commercial sector, businesses and other organisations use social data from blogs, company reviews, and social media posts to obtain consumer opinions on products, services and brands. There is general recognition that the public sector provides the most robust open data, as shown by examples such as the US Food and Drug Administration (FDA)'s 'OpenFDA' portal, which allows anyone access to publicly-available FDA data. In social terms, the US Chamber of Commerce argues, new digital applications are giving citizens access to their own personal data, therefore yielding more informed consumers. The focus here is on enabling people to be data consumers, so it does not extend to considering how they can add value to, or improve their lives by data. Furthermore, there is little attention given to how they might utilise the value of the data they produce through social media and

other data sources that harvest consumer data. Here issues of privacy are important, as discussed in Chapter Eight.

The use of open data by ordinary people is still relatively low (Cornford *et al*. 2013). In general, people without a technical background are most usually what can be termed 'data consumers'. Individuals mainly use applications that make data accessible for data consumers and rarely, if ever, use raw data. Even if individuals get involved in open data projects where they can develop new skills, they tend to rely on intermediaries such as programmers and data integrators to help them access data. The activities individuals carry out depends on which technological skills they already have. Ordinary data users include programmers, data integrators, citizens, champions, facilitators and open data advocates (Hivan and Titah 2015). These types of users often perform more than one activity, so they need to collaborate because the work is too big and complex to be done by one person. According to Hivan and Titah (*Ibid*.), there are five activities that give data value: (1) identifying data; (2) requesting data; (3) converting data; 4) programming; and 5) promoting data. Hivan and Titah's study (*Ibid*.) of open data projects in Montréal, Canada shows that these activities are identified and allocated to people through planning days which identify projects and bring citizens together to create a team. Once this stage is reached, a project has a champion designed to be a public supporter of the project (*Ibid*.).

In the context of resident-supported open data projects in cities, open data advocacy groups help to promote the applications devised for citizen's use. These groups devise open data projects that they hope will convince a city to pursue its efforts towards open data programmes and will develop an open culture within local government. A critical dimension of open data projects is assigning responsibilities to individuals to cover all the activities involved. It is important that individuals have a feeling of ownership and of actively facilitating change, because this motivates them to take part and to see the value of using data. One example is an anti-corruption hackathon that was held in Québec (Rocha 2012). Here, instead of simply complaining about corruption, individuals recruited programmers, journalists, civil society groups and other city residents to challenge the misconduct using data. The project used data to spot links and patterns between calls for public tenders and contracts awarded to specific organisations by the City of Montréal.

An important factor in enabling city residents to use open data is the design of data portals. An easy-to-use portal encourages data use and this has been identified as an important dimension in explaining the use of

data by individuals. Hivan and Titah (2015) found that Montréal's data portal is easier to use than many others. The participants in their study said that they found the portals from Boston, Toronto and New York less easy to use, and that they got lost when searching data sets. The complexity of accessing and interrogating data is a constraint on data use. In order to use data, a user must have a level of data literacy and the capacity to collect, treat, analyse, and communicate with large quantity of data. In Hivan and Titah's (*Ibid*.) study in Canada, the participants thought that city councils should be responsible for providing data literacy education for their citizens. One open data advocate in Montréal argued that data literacy is an important issue in terms of participation, because it is part of deepening an informed participation. Evidence from open data projects, such as those by Hivan and Titah in Montréal, shows that providing data literacy and easy-to-use portals are important in an open data ecosystem. Another dimension of enabling genuine participation in using data is the level of citizens' inclusiveness. Here, Berry (2008), for example, distinguishes between the image of open source information as a 'commons', and the reality, which is often that of a 'club good' enjoyed by restricted group of people who have the necessary data and technical skills to analyse and interpret the data. This point also relates to big data, as Boyd and Crawford (2012) and Carlson and Anderson (2007) note. The use of open data for social transformation towards an inclusive knowledge society therefore needs to be placed in the wider context of inclusive and interpretative communities.

Therefore, although Hivan and Titah (2015) identify ways to facilitate ordinary people to use open data, there is also a need to address the ways in which these individuals may – or may not – form into interpretive communities (Cornford *et al*. 2015). There is a strong focus by actors in the open data movement on practical aspects such as licensing, open artefacts and education, however little attention is paid to the data interpretation. As Davies and Bawa (2012) assert, openness needs to be seen as a process that is rooted in communities. They argue that this does not just depend on the open artefacts within communities, but on how they support their members' interpretive skills. Interpretative communities can be understood as 'reference groups', whose perspective provides a frame of reference for individuals within the group (Shubutani 1955). Interpretive communities also contribute towards the social production of knowledge, including formal knowledge. As Chapters Two and Three noted, it is now a well-established fact that scientific knowledge, as well as other forms of knowledge, is created by communities (Porter 1995). Communities produce a range of forms of knowledge and define the types and the qualities of the knowledge they

produce. This point leads to the consideration of interpretation – because the production of knowledge and the use of knowledge requires interpretation. The framing of interpretation occurs within communities – the cultures of communities shape the way that artefacts and symbolic goods become knowable. As Fish suggests, 'there is no single way of reading that is correct or natural, only 'ways of reading' that are extensions of community perspectives' (Fish 1980, p. 16). He argues that 'interpretation is the source of texts, facts, authors and intentions' (*Ibid*., p. 16) and we add data to this. Fish develops this idea further by saying that these entities (in which we include data), which 'were once seen as competing for the right to constrain interpretation (text, reader, author) are now all seen to be the products of interpretation' (*Ibid*., pp. 16–17). Therefore, for open data to be transformative, there need to be interpretive communities within society.

However, Cornford *et al*. (2015) argue that UK developments in open data have not reached their full potential, because insufficient attention has been focused on supporting the development of interpretive communities. Despite work by open government advocates in the UK, the current situation means that individuals tend only to be able to act as 'armchair auditors' in using open data, because of the lack of interpretive communities. Cornford *et al*. (*Ibid*.) point out that there need to be more opportunities for individuals to act effectively with data. In the first instance, this requires help from intermediary organisations which are capable of processing data and can support the interpretation of data. The learning gained from these intermediary organisations may then foster community learning, enabling individuals to gain the relevant expertise as part of an interpretive community, as open data-based participation. Cornford *et al*. (*Ibid*.) argue that support is needed to grow interpretive communities, which requires the development of an institutional context for citizen and individual engagement in the interpretation of data. This involves a move from a primary focus on openness in mechanistic terms to one that also addresses process (Joshi and Houtzager 2012). In order to use open data to transform to an open knowledge society, the social and political processes are important for ensuring that accountability within knowledge production is rooted in citizens and individuals, rather than elite governing bodies. Therefore, as Worthy (2012) argues, the development of open data within an open society framework needs to be embedded within communities – whether communities of place, interest or association – that enables the interpretation of data. Furthermore, these communities need to devise governance processes to ensure accountability for the responsible use of data. As such, the disciplinary practices examined within RECODE

as well as the stakeholder practices and inter-relationships within them, provide valuable evidence about these processes within a particular open data use case. The information gained within RECODE will provide some pathways for further development of this sector to assist in enabling these practices to develop in responsible ways.

Conclusion

The main themes of the open data movement are the value placed on 'openness' and the ways that open data can be a key driver of change, towards open science and open society as well as data-driven commercial innovation. Those in the open data movement are convinced that open data has the potential to be valuable for society, in both general and specific terms. There is a consensus about the value of open data across a wide range of social and economic life, such as open government, development and human rights, innovation and commerce. In order to mobilise open data actors within the wide remit of the open data movement, there is a need to develop protocols for open data use, easy-to-use open data portals and, as some activists urge, the development of data literacy. Realising the aspirations for open data, open society and open science is complex and it requires technological, institutional, and legal change to be embedded within the social change of achieving a knowledge society. Any transformation in developing a knowledge society using open data rests on the characteristics of the social production of knowledge. This requires citizens and organisations to have the necessary skills and time to use data in a knowledgeable way, and in ways that transform social life in a democratic and accountable way. Importantly, this means addressing the process of open data and finding ways to create and sustain interpretive communities, where there is a right to be wrong and a data welfare system. This last point is important because open data has the potential to create a participative open society – however, it can also fragment and create fragility in society. Once an openly accountable process of open data is in place, bringing open technology and interpretive communities together, then open data can play a defining role in the mobilisation of a knowledge society.

RECODE represents one of the first empirical investigations of the mobilisation of open data. The lessons learned within the project provide some information about how activists, stakeholders and reluctant members of the community are engaging with open data practices and imperatives. In addition, it provides an opportunity to evaluate policy and practice

in one sector to inform the development of these across all of the open data paradigms (open government data, open research data, big data and others). In Chapters Five to Nine, we examine some specific grand challenges – values and inter-relationships (Chapter Five), institutional practices (Chapter Six) and legal and ethical solutions (Chapter Eight) – case studies (geospatial data – Chapter Seven) and future directions (Chapter Nine) for managing open access to research data. These lessons will enable an evaluation of the progress towards, and characteristics of, an emerging knowledge society based on the availability and openness of data from governments, researchers, individuals and commercial organisations.