35

ONLINE ENVIRONMENTS AND THE FUTURE OF SOCIAL SCIENCE RESEARCH

Michael Fischer, Stephen Lyon and David Zeitlyn

INTRODUCTION

History doesn't repeat itself, but it sure rhymes.

Attributed to Mark Twain

New generations of social scientists face a different range of possibilities and prospects in their careers than many academics currently in post. The Internet and related communications technologies (IRCT) are playing a major role in these differences. The Internet has greatly impacted social scientists' practice, as well as advancing the scale of activities rendered feasible, resulting in significant changes in the kinds of research carried out and, importantly, the kinds of subject deemed 'researchable'. More important, IRCT are social infrastructures that people use to create new social phenomena, which become objects of study for social scientists.

People are using IRCT to change the world around us, creating circumstances that change quickly over such large areas, that apparently continual adaptation – technological, social and cultural – is necessary. This trend will expand apace. The opportunities for social scientists

will be driven by changes in societies and advances in our research methods, and we will perform some things better, or at least on a larger scale. We will be able to carry out hitherto unimagined activities relating to data collection, analysis and dissemination. Concurrently, many of the social and cultural forms that emerge create situations we are ill equipped to understand. We require new capabilities to enable social scientists to operationalise some well-established conceptual and terminological descriptions and understandings. We must also develop new theoretical concepts and vocabularies.

How will we deal with new kinds of social relationships? What do we do with the vast amounts of data that become available from technologically enhanced observation and participation? How will the formidable ethical issues be addressed? How do we study social and cultural phenomena that may exist for a few years, months or only weeks? How do we adapt to a dependence on 'smart' technological assistants in our research? How will we be able to disseminate our results, not just in static form but in formats that directly interact with potential users? What further technological change can we expect? Perhaps the best way to predict the short-term future (3–7 years) of the impact of IRCT on social science research is simply to look at what a minority of computer and network-savvy individuals are able to do now. Dow (1992) accurately predicted much of the development of computing in mainstream anthropology, simply by looking at what the minority were doing at the time. The contributions to this volume will serve as a model for the short-term development of IRCT-related research.

Predicting the longer term future (8–20 years) is more problematic. Today, visions and trends are evident which, if continued, will lead to identifiable future practices. However, any number of factors can interfere with current trends and derail the best-laid futurology. One can, nevertheless, still differentiate between probable, possible, improbable and (probably)

impossible applications of IRCT over the coming 20 years. Although our grasp of future history might be weak, by focusing on the development of capabilities we can get a handle on what tools and resources people (and researchers) have available to build our future.

This chapter discusses how new or expanded capabilities emergent from IRCT may contribute to changing social science research, particularly how research topics, methods and capabilities might change with increasing integration of IRCT into the daily social lives of most people in developed and developing societies. We have not limited ourselves to online research because we believe that firm distinctions between online and offline research is a present phenomenon, and that online research will rapidly become one of the many different contexts within which research is carried out – not the odd one out. However, we expect all social science research to change, for the very reasons that online research will become accepted and ordinary when online social phenomena become integrated into wider social and cultural life.

There are two broad themes: new social formations, phenomena and conditions that arise because of access to IRCT technologies; and new methods that become available to carry out social research using IRCT technologies. These two themes will, of course, co-occur and will quickly converge.

We can relate only to capabilities that may underlie research methods, not specific future methods. We discuss some of the major new capabilities which are likely and offer some examples. Similarly, we do not make specific predictions of wider social change, but rather new social capabilities. We discuss so-called virtual groups, but for the most part we shall leave predictions about specific future social and cultural development to our, and the reader's, science fiction avatars.

CHANGE AND CHANGES TO IRCT TECHNOLOGIES

Although there is a tendency to focus on technology as a material process, technology is also a process of social and cultural instantiations of ideational innovation (Fischer, 2004, 2006a) – the adaptive transformation of ideas into practice. We view technology as anything people use to extend or expand their capabilities, directly or indirectly (following Hall, 1976).

In this context, what are recognised as technologies result from ideas whose instantiation have social and cultural histories (these were successful), which in turn creates a sense of inevitability for their future. The development of material futures is never linear. Technological development and human extensions (Hall, 1976) are formed by adaptive processes. As human culture transforms the material world (Fischer, 2006a), new possibilities emerge for instantiation of our prior symbolic constructions. Core cultural ideas will also change over time, but much more slowly than how people instantiate these into the world.

Many of the visions instantiated using the Internet considerably preceded the Internet itself (see, for example, Bush, 1945). Much of the current development of IRCT instantiates broad visions (fantasies?) from the mid-twentieth century, inspired by figures such as Arthur C. Clarke, who in his fiction described global networks, networked libraries with search engines, personal videoconferencing and cell phones, and J. C. R. Licklider, whose anticipatory visions directly contributed to bringing the ARPANET (Advanced Research Projects Agency Network) to reality. However, the material forms that manifest these visions, the social and cultural formations and uses people make of the productions and interactions of these visions go well beyond what was envisaged. From a given starting point we can extrapolate *what* future capabilities there may be, but not necessarily the *forms* these will take, nor the outcomes of their manifestation and uses.

Much of what we discuss will not sound very futuristic. There is a very good reason for that because we are only looking over the next 20 years. Although people often perceive that technologies arrive and rapidly change the world around us, our experience so far is that it takes at least 15–20 years (aka the 'Fischer fifteen-year rule') for new capabilities to become pervasive following their first entry as a deliverable technology. Researchers are a bit more precocious than this, and for specialists with technical skills the period is more like 3–7 years, and specialists without technical skills up to 10 years. But for a capability to become pervasive in the research community as a whole, the period is very similar to the general public's. Much of what we discuss is partially achievable now, but is often still dependent on current and future research for continued development – so that covers the next 20 years quite well.

Although it is possible that currently unknown *fundamental* 'new technologies' may emerge over this period, it is unlikely that those would have a great impact for at least 10 or 20 years afterwards. For example, microcomputer technology was first delivered by Intel as a commercial technology in 1968, and gained mass acceptance in the form of microcomputers in the period between 1983 and 1985. Email was first introduced on the ARPANET around 1972–3, but did not achieve mass acceptance in universities until around 1988. Telnet (for interactive sessions between networked computers) was also introduced in 1972–3 and FTP (File Transfer Protocol, for file transfers between networked computers) in 1973. The 'web' was introduced in 1991 between a few institutions, expanded slowly during 1993–4, and began to become a phenomenon from mid-1994 (after the release of Windows 95) – nearly 22 years after FTP (whose functionality it incorporates) and 18 years after the first public online information services (Leiner *et al*. 2012 [2003]).

Any increase in IRCT-mediated ('online') social relations will result in social change by definition. A principal topic of this volume is just how we, as social scientists, should go about the study of these relationships. For instance, some researchers have been attracted to online research because of the appearance of new online communities. Others have been attracted to the use of online panels and surveys to study more traditional social institutions and formations, and many researchers are simply grappling with the impact of IRCT on what they would consider to be more conventional research settings.

The use of technology in social science research is hardly new (or uncontested). But IRCT supports many new opportunities and capabilities for data collection and documentation, theory and analysis. Aspects of the research process that IRCT can most greatly impact are:

Communication – the capacity to gather, disseminate and exchange information. This includes data collection, whether through direct contact with people or by sensors (cameras, global positioning systems (GPS), heart-rate monitors and environmental sensors), collaboration with researched colleagues or research colleagues and dissemination of the outcomes of research.

Representation – the capacity to describe, model and visualise information: how information is aggregated, visualised, described, modelled, transcribed, presented, transformed, reduced, expanded and interrelated.

Storage – the capacity to retain and retrieve information: the form, medium and availability of retained information (most often representations).

IRCT greatly enhances the scope and integration of each of these processes in research; communications is no longer an end-point after the fact, but an integral part of the computational environment. Code, processes and data can be distributed across the network, greatly expanding not only the capacity of researchers to exchange and share resources, but also transforming how research is done, its replicability and the production of sustainable outcomes.

COMMUNICATIONS

Communicating complex symbolic messages did not, of course, begin with the Internet. Generative language development and then writing, respectively, made new kinds of social organisation possible, although the strongest forms of this claim have been questioned (see Goody and Watt 1963). Eisenstein (1979) posited similar radical changes following the printing press (see also Zeitlyn, 2001).

The advent of telegraphy, telephony, radio, photography, film and television each had profound impacts on how people were able to record, transmit and use information that cannot be subsumed within the capabilities originating with language, writing and printing. Each technological development enables new means for forming and maintaining social relationships, while rendering some types of social relationship less critical or obsolete. Internet communication via email, conferencing and collaborative web applications transform the ways in which social scientists can exchange information and develop friendships and collaborations. The gamut of FTP and resulting services enables sharing immensely large distributed datasets of disparate data types with relatively low cost and effort.

The current rise of mobile Internet platforms, such as phones and tablets (see Silver and Bulloch, this volume), has radically transformed the concept of locale. As video-based communications has spread to phones and tablets, most researchers have participated in a video conference at some stage of their research. Although there is still a vague scepticism about the ability of such formats to genuinely replace more conventional forms of meetings, this scepticism is rapidly receding. Replacing meetings was the original trajectory for video-conferencing, but improvements in video presence technologies in conjunction with mobile devices have increased the frequency of communications, irrespective of the impact on physical meetings.

One of the obvious growth areas in Internet communications is transmission of this real time 'presence' data, including audio, video, live camera feeds, physiological measurements such as heart rate, and geographical location. The main trend of developing capabilities over the next two decades in research communication will be increasing pervasiveness in exchanging expanded indices of presence. Presence is what we individually bring to a situation and context. Communicating presence brings more of ourselves and the others we interact with into a common context. The telephone was a great stride in presence, and found its way into the research process, sometimes controversially (at least where sampling was an issue). Increasingly 'presence' will refer to our ability to exert influence or be influenced, physically or otherwise, over a communications link.

Improvements in sensors and actuators will enable transmission not only of sound and image but also of heat, odour, taste and surface texture. Transmissions will be not just as digital representations but increasing with the capacity to materially reproduce these at all networked locations. We will meet in simulated environments for demonstrations, meetings, data collection or processing using simulated representations of ourselves and others in simulated space transposed over a shared composite locale. Interactions will not be limited to the simulated space; we will link actions that we and others take in our local locale to reactions in the composite locale. The effective transmission of material objects over communications channels will be commonplace because instructions will be sent to devices that manufacture objects (perhaps like an elaborate 3D printer or using 'smart' materials that reassemble themselves into requested forms).

What is likely to transform the way social scientists carry out their work is the pervasiveness and the complexity of the presence-focused communication. Current mobile communication devices have substantial capacity for complex communication including file

transfer, video and audio in synchronous and asynchronous modes. Moreover, much of the communication does not happen between two people directly, but with some form of software agent acting as mediator, directly engaging in the communication. At the moment we can see this in electronic calendars, Amazon-style user-focused pages recommending further purchases based on previous browsing history, or social computing sites such as Facebook, Twitter or Reddit where software agents create personalised resources or viewpoints. Software will continue this trend in simulating people to the extent that routine conversations may well be with (or between) software agents that brief their 'operator' later. The only way humans may be able to differentiate some communication between software agents or people is the inefficiency and delayed response time of the person.

With respect to research practices, we anticipate three relevant types of change:Changes to the profile of potential collaborative partners;Changes in the ways certain kinds of 'field' research may be conducted;Changes in the ways in which the mundane aspects of being a member of an institution are acted out.

Network services already make possible geographically distributed teams of researchers who coordinate their efforts and effectively create something akin to research centres without a physical location. In 1995, Zeitlyn created a Virtual Institute of Mambila Studies, which brings together resources relevant to the international pool of Mambila specialists. More recently, many projects in the Social Sciences, such as Kinsource[i] and Complex Social Science Gateway,[ii] have emerged, involving many individuals and organisations distributed globally to construct, use and collaborate in specialised research areas, not only with shared data, but also shared resources and tools for analysis that leverage shared data. Organisations such as the Human Relations Area Files (HRAF)[iii] are refactoring their current web application into a suite of software services researchers use to greatly customise access to

HRAF data, the procedures applied to data and the creation of sharable documents containing outcomes of searches or analyses, all of which utilise network communications to reference a common set of data on the Internet.

Changes to siting the 'field' are underway. Webcams constitute a legitimate area of study for the social sciences.[iv] The capacity for streaming 'presence' data changes how primary field data can be collected, disseminated and made available for secondary research. Social media will result in more and more 'traces' of people's presence. Short-term field research combined with judicious use of networked presence data in partnership with local academics and informants is potentially a means for collecting ethnographic data and increasing the reliability of those data.

To summarise, much of the 'future' of pervasive communication is in fact the present! Little new technology is required to achieve the ubiquitous disparate communications context we believe is emerging. However, new technological developments will enhance many aspects of this communication and widen the range of people using it. We can predict some outcomes:

1. Collaboration will rely on pervasive multi-format interactions, all of which are possible today, but which will be simpler, more integrated and more robust.

2. As such communication becomes more pervasive, the objections about impersonality or partiality will recede. In other words, people will develop new ways of inferring closeness, intimacy and trust through online interaction.

3. Individuals will change their assumptions about privacy and trust, as currently suggested by subdued reactions to increasingly regular cases of personal data being lost, stolen or leaked from financial organizations, insurers and government, which are regarded more as inconveniences than major scandals.

4. Pervasive online communication, like simple email and multimedia presentation software before it, will become part of the baseline set of software tools that all social scientists will be assumed to have mastered.

REPRESENTATION

When collecting data and documenting human practices, institutions, languages, societies and cultures, social science researchers directly incorporate new technologies of representation in a primary sense, and also data derived from what people create using the technologies (new and old) at their disposal. Data is derived from and is represented by fieldnotes, sketches, transcription, photography, telephones, radio, audio recording, film and video, and – increasingly common these days – interactive media distributed over the Internet (Macfarlane, 1987; Farnell, 1995; Biella, 1997; Fischer and Zeitlyn, 1999).

Researchers are familiar with recording aural and visual data as part of data collection. These recordings can be used reflexively in the field to elicit detailed descriptions, to interpret and to disseminate knowledge. The advent of hypertext expands the capability to interrelate components of both data sources and data representations, with the addition of links between segments of different media, allowing researchers to record knowledge about the interoperation of the people, processes and objects depicted by the media, both their own and knowledge elicited from their local research collaborators on the ground (Biella, 2004; Ruby, 2005). This capability has, however, been little used by mainstream researchers.

Computer representations have generally been considered by most people as virtual objects – abstract representations of real things. Increasingly, computer representations are achieving first-class object status, where people can manipulate and exchange these as they would 'real' objects. Initially for video game players, and more recently for users of mobile technology such as the iPhone, configurable objects are increasingly common in people's lives, mediating interactions between people, and thus becoming as much objects of social

research as any other human artefact. Inexpensive hand-held 3D scanners are becoming available on phones or tablets, producing hybrid images that integrate a 3D mesh representation of an object rendered on the surface with photographic data. These are objects that can be further manipulated with computer-based tools, imported into new scenes and material copies reproduced on a 3D printer. In conjunction with development in 3D capture and display technologies, such 3D objects will increasingly replace 2D digital photography and video for research. Rather than simple recordings of light, a recorded event will have discrete objects interacting with each other, objects with persistence in the recording that can be associated with further data and identified in other recordings.

The development of mobile computing platforms and improvements in authoring complex interactive media creates the capability for recording physical interaction with embedded media objects available in the field (Zeitlyn and Fischer, 1999; Bagg *et al.*, 2006). Phones and tablets already have software for single platform capture, editing and display of media, and mobile platforms will replace conventional cameras, computers and displays for most researchers, as well as the general population. Developments in projective and perceptual displays will make mobile platforms more mobile, in the form of watches, rings, pendants and badges. Widespread subcutaneous cyborg modifications beyond medical applications, where hardware is embedded directly within the body, is likely to remain mostly a minority activity over the next two decades, although we can anticipate governments and corporations to promote 'ID chipping' of people and parents chipping their children.

The availability of embedded computers and computer sensors will greatly extend capability. In 2016, tiny computers with speed and storage roughly comparable to desktop computers of just a few years ago are commodities. These are miniaturised to a size somewhat smaller than a fingernail, very inexpensive and able to operate for substantial

periods on small power cells. These will use similarly miniaturised sensors that can measure and record many details of a person's interaction with their environment and with other people, including proximity, motion, acceleration, rotation, skin temperature, brain and nerve activity, blood chemistry and anything else that can be measured.

For example, presently researchers, tourists, and nearly anyone with a phone are using GPS technology in conjunction with digital photography and video to add spatial and temporal location to the mix of relationships that are recorded with the image (Fischer, 2003; also Happel, 2005). The research day, week or season can be played back temporally and spatially (say on a map), evoking recorded media, notes and other time-stamped data that is associated with the researcher's presence (Fischer, 2006b).

Similarly, social networking is beginning to draw on sensor readings, for example GPS functionality in photo tags can invoke Google Maps to display where the photo was taken, and Nike+ offers a running shoe that logs information regarding the run to an iPod Nano and then uploads data to the Nike+ website [v] where runners can compare runs. Social apps such as foresquare.com alert users when they are in the proximity of friends or other users meeting a certain profile.

In other words, the trend is to increase our capacity to record much more of the research context and process, and this greatly expands the kinds of data we have accessible to us, including sensor data recorded by potential research subjects on their own initiative. Multi-megapixel photography and HD Video, combined with new, cheap 360×180-degree lenses, already make it possible to visually record a complete scene, not just an aperture of a few degrees.

All this will, of course, create new issues for how to represent and use this staggering array of data. Conventional methods, such as statistical summarisation of particular views of the data, will of course continue to be used. But we will be increasingly driven to disaggregated designs, where we build layers of abstraction and aggregation over the dataset while retaining links to the underlying data. Some data will be real time streams, constantly generated by the activity of potential research subjects. If not 'on-line', data will increasingly be 'on-tap'. Research design will generally transcend towards disaggregation and data reuse.

Embedded systems can control actuators that translate data into effects in the world. Common actuators currently mostly produce movement, sound, light and heat, but texture mapping and odour synthesis have been demonstrated, and in principle any sense can be reproduced individually. The opportunities for aggregating these into research data representations are, as the 1970s microcomputer sales slogan stated, only limited by our imagination. Certainly a range of new research based on controlled experiences is likely, as well as the production of 'identikit' data instruments where people create experiences for the benefit of the researcher as data.

There will be a very strong technological push over the next two decades outside the social science community for development of multi-sense sensors and actuators, driven by a major industry theme often referred to as the Internet of Things (Madakam *et al.*, 2015) or IoT. The broad conception is literally to put everything in the world directly online, by either observing it, or attaching sensors and actuators to it, all interfaced to the Internet. These will range from household appliances that report and track their contents to the deployment of billions of small sensors into public and private environments, creating smart environments that track any interaction and make this data available on the Internet, as well as perhaps being able to display personal public service information (or personally focused

advertisements) on the lawn of a public park. Social scientists have a range of opportunities and responsibilities over this period, if nothing else to help ensure that this does not result in a surveillance and control system that far exceeds the worst nightmares of George Orwell. But these plans almost guarantee that, even if the dream (or nightmare) of the IoT fails for some technical or social reason, there will be an unprecedented amount of data regarding people and their interactions with each other and the environment around them.

There are two basic issues that emerge in relating these capabilities to research methods. The ethical dimensions of research on this scale, which depends on near or real time information relating directly to individuals, are vast. But at present this level of detail is largely irrelevant to our research questions and research methods, and in many ways, beyond our present conceptual capacity.

There are, however, connections with existing research methodologies. Ethnographic studies, although usually on a smaller scale, have encompassed much larger communities by using a combination of immersive observation in sub-groups, whilst evaluating the results of immersive observation through sampling the larger population (Moody and White, 2003; also Fischer, 2006c). Mass observation studies have made sense of the records of thousands of people's self-observation. Larson and Csikszentmihalyi (1983; Csikszentmihalyi, 1991) – introduced 'beeper' technology to ground and contextualise the interactions of large research populations, with participants reporting activities under way when the beeper sounded. Each of these techniques seeks to impart meaning to the behaviours that can be observed.

At first blush it appears that all we get from the capability to access large sets of detailed data is a lot of behavioural data, with no meanings associated with that data. But because it is all disaggregated data, there are opportunities to do a great deal more. In the early days of research using satellite imagery a similar situation prevailed. There were many measurements

of different aspects of an area, but researchers could not assess much more than what the measurements themselves entailed: how much light of different frequencies was reflected. To use this data for environmental research, research was done to examine the areas the images represented, producing baseline data on physical topography, plant cover, buildings, crops, fields, bodies of water, vehicles and other objects, which were then related to the imagery. The outcome of this process made it possible to identify similar 'ground-truth' areas in new locations.

What will be needed is the development of 'proofing' subsets of the behavioural data, in order that findings from the 'proofed' data can be extended to the larger set of observations. Methods for this purpose are under development and are included broadly within the relatively new research activity of data mining (see Ayelet-Tsabari *et al.*, this volume). Data mining depends on relating patterns in disaggregated data streams to knowledge (and sometimes guesses) about the processes that produce that data. Rather than a return to pure behaviourism for all social scientists, we can therefore use the behavioural outcomes of ideationally driven processes as indices for identifying the likely presence of similar processes elsewhere. Thus, data mining can, in limited circumstances, replicate emic driven processing by people.

This methodology is related to many present social science research perspectives. Some of us carry out small-scale ethnographic studies, or focus groups, or do sample surveys of some fragment of a population. We attempt to identify the social processes at work in these studies. We then attempt to generalise the results, based on ethnic or cultural group, social group, educational group, language group, etc. The principal difference here is that we are directly relating the patterns we observe and have 'proofed' to the larger population, not just through a few well-studied proxies.

New methods and means of representation and visualisation developed to support e-Science (Fielding, 2003; Fielding and Macintyre, 2006), multi-agent based simulation, shared network tools, and the Internet of Things (Madakam *et al.*, 2015) will increase our capacity to work with multiple views of the disaggregated data (Bainbridge, 2007), enabling multiple research designs to be instantiated during, or even after, the data collection, the use of hybrid designs such as interactive dynamic statistical sampling, and composite representations that are 'layered' so that the original data is always available regardless of the level of abstraction (Fischer, 1998).

If considerable ethical issues can be resolved, with sufficient resources it becomes possible to track the movements and interactions, visual and aural context and the 'presence' data of an entire population.

STORAGE

Recent developments in 'intelligent' machine data storage have produced conceptual tools that are certain to have an impact on the kinds of research social scientists are not only able to imagine, but indeed will be required to conduct. The present model of storage has been to associate particular bits of information with particular places. The advent of Internet search engines demonstrates that this model has seen its day. There is simply too much information in too many places to organise using a simple set of addresses or locations.

One possibility is to access information based on its content (semantically) rather than its location. The idea of semantic or associative storage has a long history, in fact it goes back to the visionary paper by Bush (1945). It was founded in one of the earliest programming languages, Lisp in 1958 (see McCarthy, 1979), and has appeared more recently in the Semantic Web (Fensel *et al.*, 2002). The semantic storage concept goes beyond matching

content, as with keywords or classifiers, but rather depends on a model of 'understanding' the content and entailments of the content in different contexts.

Semantic storage systems enable software to infer meaning from data and relationships between data. There have been a number of increasingly sophisticated partial solutions to the problems, working around the fact that machines do not think as humans do; that is to say, that although a human with a reasonable search engine is capable of identifying related information across a range of websites, a machine is greatly handicapped by the ways in which such data is currently stored, largely because as yet we have not been able to model how we understand the content. Another approach, which underlies much of what makes search engines such as Google work well for some applications, is effectively based on data mining – the choices that people make after they do a search (what they clicked on) is recorded. Future searches are ranked on how close these are to past searches, and tend to 'promote' popular choices from those searches. Over time, each search is augmented by earlier searchers' choices.

Most current solutions involve adding different kinds of metadata (what machines use to infer relationships) to the content, and this has made it possible to produce prototypical versions of a Semantic Web, in which a range of inferences may be generated automatically. At present there are limitations imposed by the absence of such metadata in most web repositories, as well as scalability problems (Owens, 2005). The scalability issue is sure to be resolved, but the absence of pervasive metadata on the web is not as easily addressed. Data formats such as RDF (Resource Description Framework) and OWL (Ontology Web Language) are based on describing data relationships using terms and relationships in subject 'ontologies' in order that the researcher draws 'semantic' inferences from data sets stored in this format based on models defined by a researcher or standardised models supported by the

research community. These are simply not, at present, designed with most social scientists (or many other categories of people for that matter) in mind. Part of the problem is the amount of specialist labour required to classify each online resource to fit the classification scheme that permits inference to take place (Brent, this volume, highlights this issue). This will change in part through better integration of social science knowledge of how people organise complex data. Kinship terminologies, for example, offer a very simple, but yet very robust algebraic mechanism for ordering relationships of extremely large numbers of individual people (Read, Fischer and Lehmann 2014). Other sorts of indigenous systems used to order the natural world share similar properties of simplicity, with impressive scalability, which are, at present, arguably limited by aspects of human cognition other than the inference systems themselves. Greater inclusion of natural or evolved human systems of inferring relationships, we expect, will enhance the capacity of human users to make ever greater use of the vast array of complex data available.

Indeed, we see evidence that such mechanisms for ordering relationships are already being successfully implemented in social networking sites in two ways. First, the sites ask users to classify friends according to a set of criteria, which will then enable relationships between friends of friends to emerge; second, friends in common automatically get highlighted, which enables a certain measure of the coherence of a given set of networks (see Hogan, this volume). Similarly, sites such as Flickr and Digg serve as an online folksonomy, where users create their own labels or 'tags' for images and web resources.

Folksonomy sites, where people are increasingly tagging most of what they create themselves in their own terms, combined with our own research on how people organise and use knowledge, should provide rich data for social science research AND have applications to

creating the Semantic Web. At the end of this process we can look to having intelligent 'assistants' to help us identify and analyse data, rather than simple workstations on our desks.

All of these content-based approaches highlight a serious upcoming dilemma for social scientists. All of these depend on making judgements regarding content, effectively aggregating the data based on particular biases or goals. The extent to which researchers are isolated from the criteria underlying these judgements represents the extent to which they are isolated from the fully disaggregated data. However, there will be too much data with too much complexity for most researchers to work with it directly. We will have to wait to see precisely how research evolves to resolve, or at least limit, the impact of this approach. Solutions will probably depend on various kinds of triangulation, development of researcher controls over the process, and new understandings of broader more holistic data environments within which many of these problems may simply be rendered irrelevant.

SOCIAL CHANGE

The immediate basis for discussing possible future social change is change in the period from 1990 to 2015, much of which is discussed in this collection. We have argued that the major driver of social change from IRCT is a trend towards pervasive, and even ubiquitous, communication. Since 1990 email has developed from a niche mode of communication for academics to a mainstream medium worldwide. This trend is not confined to the Internet. Seemingly, regardless of economic circumstances, mobile phones, once mainly a source of irritation in restaurants and trains, are a possession of the majority of people in most nations. Access to the Internet has changed from episodic connections using simple modems to pervasive connections via mobile or landline broadband, and increasingly using high-speed fibre-based or high-speed mobile connections, with a strong trend towards 'always-on' mobile connections and applications.

In the developed world we already have the capacity for pervasive communication. We can phone, email, instant message (IM) or text most of our social partners at any time, as can they. We interact often on social Internet sites. Our ways of interacting with each other are adapting rapidly, particularly among the young, whose opportunities for physical contact are becoming increasingly restricted. There are imbalances based on relative income, but surprisingly this absolute gap, at least in terms of being connected at all, has diminished rather than enlarged. This is true for nations with emerging economies as well, where some of the poorest nations on Earth have 70 per cent or more individual connectivity at some level for mobile networks.

Currently, communication is dominated by written and spoken language and, to a more limited extent, images, still or animated. Although the episodic period is very much reduced, there remains a socially imposed periodicity on communication. While the generations born prior to 1975 tend to regard privacy as an important element of their lives, those born since 1985 are much more apt to regard any aspect of their lives as public, though in their control. The rise of social sites in the period following 2002 has resulted in vast amounts of information about day-to-day private life being published on the Internet. In 1999, Scott McNealy, then CEO of Sun Microsystems, commented, 'You have zero privacy anyway. Get over it' (from Sprenger, 1999). If the ethos of the 1960s was reflected in Andy Warhol's suggestion that everyone could have 'fifteen minutes of fame', by 2030 it will likely be radical to offer people 'fifteen minutes of anonymity.''.

Since the appearance of the first webcam in 1993, hundreds of people have published their lives on the Internet, and hundreds of millions regularly provide day-to-day details, photographs and videos. Increasingly, individuals will use pervasive wireless networks to broadcast their day in progress, at least to what they perceive to be their social network.

Conventions of management of image will evolve with both transmission and access to this information. It will not be a 'raw' transparent record, but another tool in presentation of self and of group, perhaps even designed to 'edit' the public record available otherwise.

The use of CCTV has expanded greatly in the period up to 2015, and is likely to continue. Countries like the UK have vast numbers of cameras covering city centres, shopping outlets, and – increasingly – residential streets. Plans to 'chip' vehicles, together with sensors in the roads, will track movements. Mobile phones can be tracked using either triangulation to transmitters or, increasingly, embedded GPS. Individuals are placing GPS trackers in their vehicles (and on their children) that can 'phone home' coordinates when the car starts, leaves a specified zone, or operates at high speed, and can be phoned to covertly listen in.

It is likely that over the next two decades more and more use of cameras, 'smart' ID cards, chipped pets, chipped children, environmental sensors in smart environments and the Internet of things, together with peoples' own choices and interactions, will be accessible online, probably to a large extent publicly, so that 'privacy' groups may force public access as the only solution to protecting people from specialist government and commercial surveillance, transforming a threat to a resource that will modify social relations. With respect to online research and social science research in general, more and more information will be available to us, and our potential research subjects will themselves be using this information as a part of forming their own lives, and thus of the meanings that they manage. Increasingly these relationships will be conducted and managed online.

VIRTUAL COMMUNITIES

One outcome that will emerge from this increasing capacity to 'know' people from their online presence is a great realignment of how people manage social relationships. Robin Dunbar argues that individual people can efficiently manage social relationships based on

personal knowledge in relatively small numbers, about 150–200 in total (Dunbar, 1993). Although we might want to quibble on the actual quantity, numbers of at most a few hundreds are consistent with most studies of personal networks and ethnographic accounts (de Ruiter *et al.*, 2011). People faced with this much information, on so many people, could be expected to either substitute 'virtual' relationships for locally situated relationships, or to develop culturally acceptable technological aids to managing more relationships, as has been the long-standing practice of sales folk, account managers and ethnographers.

Castells (1996, 2001) refers to real virtuality, as opposed to virtual reality; by that he means the virtual space which becomes as real and integral to people's lives as more traditionally recognisable realities. Cyber communities are cropping up and creating ways to fill in the gaps of online sociality and render it increasingly 'real', with increasingly ambitious achievements in the 'real' world. Initially this was largely of interest to social scientists interested in studying themed groups or marginalised groups that for one reason or another found it difficult or impossible to be more open in their community activities, but the techniques for people to overcome the impersonal nature of socialising on social computing sites are emerging and easier to implement and interpret.

Online sociality has developed over the past 40 years from technically apt special purpose groups, such as those underlying the forums of HumanNets on the ARPANET, to whole new forms of sociality; the groups within Open Source, who have redefined concepts for intellectual property, groups that have contributed to political change such as MoveOn.org and the movements that emerged in the Arab Spring. Groups have formed around prior social relations, such as Facebook or LinkedIn, as well as groups that spread information, such as Twitter or Huffpost, and countless groups that organise around themes (such as space travel,

boating or writing) who communicate largely through contributions to building a joint resource with limited person to person communication (Applin, 2014).

Developments such as these support the view that most of the present focus on 'virtual relationships' should, following Castell's lead, be seen as a variation of 'actual' social relationships. These relationships are not virtual, but simply based on new forms of reciprocation or exchange, and indeed it is likely that such social relationships in the future will be based on more 'real' information than at present. In any case the boot-strapping processes for children and young people transforming the 'virtual community' into 'community' are already well established.

TEMPORARY COMMUNITIES

Temporary communities offer a number of opportunities for social scientists. When people come together for a common cause, motivated by interests which have, to some extent, built-in expiry dates, it becomes possible to observe conscious community-building techniques. Many of these will fail because the people involved have never seriously tried to understand what makes communities remain cohesive through differences of opinion, disagreements about resource allocation and the host of other incidents that arise and cause people to decide they would be better off either with another group or on their own. Primate and hunter–gatherer populations demonstrate the propensity of small groups to have very fluid group composition and to break up and rejoin frequently. With sedentarisation comes the need for more complex mechanisms for conflict resolution and negotiation. Interestingly, the kinds of special-interest community made possible by IRCT may need far simpler and less robust conflict–resolution mechanisms because the scope of interaction is highly restricted. Moveon.org had effectively developed an online movement more or less in opposition to George Bush and the War on Terror. It is almost inconceivable that all the members of

Moveon.org would cooperate well in face-to-face settings, and even less likely that they would agree on all the major issues in foreign policy confronting the US.

Nevertheless, in a sense such a movement is evidence of IRCT's ability to foster temporary communities around restricted sets of issues. The communities need not be tested in the way residential neighbourhoods might be because one will never be confronted with the reality that one's community fellows in fact are selfish, or xenophobic on some issues, or sexist or racist in some ways. To some extent, the members may imbue other members with agreeable characteristics, using the logic that if someone was against the War on Terror, or did not care for George Bush as President of the US, then he or she must also agree with me on X, Y or Z. Using such logic, it becomes possible to create very powerful online communities with limited capacity for longevity. When over time conditions underlying the original formation of the group are resolved, then many such movements will disappear as well. Much as the war protests against Vietnam created odd bedfellows in the US, so too can opposition to global events create unusual coalitions of individuals. What makes these interesting, and possibly the result of a kind of IRCT revolution, is their pervasively distributed locality. Apart from the fact that the most widespread of such temporary communities, for the moment, use English as their language of communication, they bring together the IRCT-savvy individuals from literally around the world. We expect that such temporary communities will rise and fall with increasing rapidity, and that one of the areas of social science investigation will be when and where such communities arise and why. Clearly not all the actions of global capitalism have provoked successful temporary resistance communities, despite the fact that some individuals will almost certainly try, and so it will be the task of social scientists to identify possible causes for success or failure of such groups.

CHANGE IN ETHICAL STANDARDS

Social scientists' awareness of ethical standards greatly increased over the latter half of the twentieth century, and over the next decade or so it is likely that ethical attitudes, and thus ethical standards, will change substantially. Eynon *et al.* (this volume) discuss many relevant ethical and legal issues that can be extrapolated into the future.

It is already clear that social scientists' attitudes towards privacy are lagging well behind public standards, while the societies around us are tolerating, if not promoting, ever-escalating, hair-raising contexts as entertainment. It is also clear that informed consent cannot be obtained for most webcam streams or satellite imagery. Streams of 'presence' data from 'smart environments' in the future will likely be similar. Is it ethical to do research based on such public resources? If we decide it is, is it still so if we commission the camera or smart environment?

As attitudes in our culture and society shift and privacy is redefined, we can expect our own ethical attitudes to change, and with these ethical standards of research. We are each, in our respective research communities, going to have to arrive at decisions about what we can and cannot use ethically in our research.

COMPLEXITY

It is clear that those social scientists who take up the challenge of dipping into this vast vortex of data will require methods that are different from the norm today. The foundation for appropriate methods is already being developed by social scientists, including the contributors to this volume, and others who are adapting research methods from the physical sciences trading under the 'Complex Systems' label (for example, Human Complex Systems, UCLA; Santa Fe Institute, Complex Social Science Gateway, UC Irvine). Basically the complex systems approach represents a union between small-group or individual studies producing disaggregated research, and large aggregated studies that have typically depended

on mathematical summarisation. The basic idea underlying research involving complex systems is that most social phenomena 'emerge' from the interaction of individuals and their contexts, which are ever-changing because of the actions of individuals and the emergent nature of social phenomena.

The complex systems approach crosses most of the traditional divides that have developed in the social sciences: it is both reductionist and non-reductionist, aggregated and disaggregated, symbolic and material, macro and micro, formal and informal. The area is also fiercely interdisciplinary and multi-disciplinary. Research methods depend on collecting data and representing explicitly and individually all the agents in a process, usually heterogeneous agents who all have their individual properties as well as their discrete representation. Agents may be represented by a few heterogeneous features or variables, or with a great deal of fidelity. Examples of this approach in social science have included studies of crowd behaviour, drug addiction (Agar, 2005), pastoral nomads (Kuznar, 2005; White, 2004), agricultural change (Fischer, 2002), and social change in institutions (Fischer, 2006c). Even where there are small numbers of heterogeneous agents, the complexity of creating models where the phenomena under study can emerge generally requires computing support. Larger models challenge the capacity of high performance computing, requiring facilities similar to those required by astronomers who model galaxies and physicists who model entire atmospheres, molecule by molecule.

Although the study of Human Complex Systems under the complexity/emergence paradigm is still in its early days, this would appear to be an appropriate way to utilise the greater volume of data we anticipate within the socially more complex formations we expect to form. However, the techniques being developed, the cyber-infrastructure that will be developed to accommodate this research and the issues that will emerge from this research

should supplement, not replace, existing approaches to research. Nevertheless, even 'conventional' research methods must be adapted to the scope of data used, matching small case results to large-scale databases, incorporating advances in theory that emerge, and determining how to adaptively use new techniques such as agent-based modelling and data mining, which also represent viable approaches to working with large amounts of continuous data (see Elsenbroich, this volume).

CONCLUSION

On the one hand, much of what we have 'predicted' is in fact already possible and already being done – but only in small numbers and by a relatively computer-savvy elite/minority. But software tools will become easier to use and will no longer be the exclusive domain of a technological elite. The network society is an increasingly pervasive reality that social scientists will not be able (or want) to ignore. The information society is either around the corner, or we are already in the middle of it. Perhaps we will know which in 10 years' time; but we can be certain that whether it is here now or just imminent, the world has changed from 20 years ago. In 1970, Alvin Toffler's *Future Shock* (1970) articulated what life-as-normal was to be for all of us from now on. It is no longer just the baby-boomers who are lost in the world they have found as adults – it would appear that every generation is doomed to look back on their childhood world and wonder where it went. The flow of information and capital has introduced a greater demand for resilience and flexibility and a willingness, or at least an ability, to re-form oneself and one's community attachments based on a shifting set of contingencies. Although the likes of Manuel Castells (1996) and Frank Webster (1995) perceptively recognised the broad strokes of such a transformation in the 1990s (and even, to a lesser extent Daniel Bell in his post-industrial society formulation of the early 1970s), it remains the task of social scientists to put the empirical flesh on the bones of such grand social theory and to identify specific mechanisms for coping with such a shifting and

uncertain dynamism at the level of real individuals and real communities, either virtually real or really real.

REFERENCES

Agar, Michael (2005) 'Agents in living color: towards emic agent-based models', *Journal of Artificial Societies and Social Simulations*, 8 (1). Available at http://jasss.soc.surrey.ac.uk/8/1/4.html (accessed 15 November 2015).

Applin, Sally (2016)   *In and Around Silicon Valley: Making, Being, Branding*. PhD dissertation, University of Kent, Canterbury,

Bagg, Janet, Fischer, Michael and Zeitlyn, David (2006) 'ImageInterviewer'. In Michael Fischer and David Zeitlyn (eds.), *AnthroMethods*. Canterbury, UK: CSAC Monographs.

Bainbridge, William Sims (2007) 'The scientific research potential of virtual worlds', *Science*, 317 (5837): 472–6.

Biella, Peter (1997) *Yanomamo Interactive: Understanding the Ax Fight on CD-ROM* (Case Studies in Cultural Anthropology Multimedia). Belmont, CA: Wadsworth Publishing.

Biella, Peter (2004) *Maasai Interactive* (CDROM). Belmont, CA: Wadsworth Publishing.

Bush, Vannevar (1945) 'As we may think'. *Atlantic Monthly*, July 1945, 101–8. Available at http://www.theatlantic.com/magazine/archive/1945/07/as-we-may-think/303881/ (accessed 16 August 2016).

Castells, Manuel (1996) *The Rise of the Network Society*. Oxford: Blackwell.

Castells, Manuel (2001) *The Internet Galaxy: Reflections on the Internet, Business, and Society*. Oxford: Oxford University Press.

Csikszentmihalyi, Mihaly (1991) *Flow: The Psychology of Optimal Experience*. New York, NY: Harper Collins.

de Ruiter, Jan, Weston, Gavin and Lyon, Stephen M. (2011) 'Dunbar's number: group size and brain physiology in humans re-examined', *American Anthropologist,* 113(4): 557–68.

Dow, James (1992) 'New direction for computer applications for anthropologists'. In Margaret, S. Boone and John J. Wood (eds.), *Computing Applications for Anthropologists*. Belmont, CA: Wadsworth, pp. 267–82.

Dunbar, Robin I.M. (1993) 'Co-evolution of neocortical size, group size and language in humans', *Behavioral and Brain Sciences*, 16 (4): 681–735.

Eisenstein, Elizabeth L. (1979/1997) *The Printing Press as an Agent of Change: Communications and Cultural Transformations in Early-Modern Europe*. Cambridge: Cambridge University Press.

Farnell, Brenda (1995) *Wiyuta: Assiniboine Storytelling with Signs* (CD-ROM). Austin: University of Texas Press.

Fensel, Dieter, Wahlster, Wolfgang, Lieberman, Henry and Hendler, James (2002) *Spinning the Semantic Web: Bringing the World Wide Web to Its Full Potential*. Cambridge, MA: MIT Press.

Fielding, Nigel (2003) 'Qualitative research and e-social science: appraising the potential', Swindon, UK: Economic and Social Research Council.

Fielding, Nigel and Macintyre, Maria (2006) 'Access grid nodes in field research', *Sociological Research Online*, 11 (2). Available at http://www.socresonline.org.uk/11/2/fielding.html (accessed 15 November 2015).

Fischer, Michael (1998) 'Counting things and interpreting ideas: anthropological conventions in the use of "hard" versus "soft" models'. In M. Fisher (ed.), *Postmodem Applications to Natural Resources Development.* Canterbury, UK: CSAC Monographs, pp. 43–77. Available at http://lucy.ukc.ac.uk/PostModern/ (accessed 15 November 2015).

Fischer, Michael (2002). Indigenous knowledge and Expert Knowledge in Development. In ed. Paul Silatoe and Alan Bicker (eds.), *The Contribution of Indigenous Knowledge to Economic Development.* London: Harwood.

Fischer, Michael (2003) 'Research note on linking GPS co-ordinates and visual media'. In Michael Fischer and David Zeitlyn (eds.), *AnthroMethods*. Canterbury, UK: CSAC Monographs.

Fischer, Michael (2004) 'Integrating anthropological approaches to the study of culture: the "hard" and the "soft"', *Cybernetics and Systems*, 35 (2/3): 147–62.

Fischer, Michael (2006a) 'Cultural agents: a community of minds'. In Oguz Dikenelli, Marie-Pierre Gleizes and Alessandro Ricci (eds.), *Engineering Societies in the Agents World VI. Lecture Notes in Computing Science*. Berlin: SpringerVerlag, pp. 259–74.

Fischer, Michael (2006b) 'Configuring anthropology', *Social Science Computing Review*, 24 (1): 3–14.

Fischer, Michael (2006c) 'The ideation and instantiation of arranging marriage within an urban community in Pakistan, 1982–2000', *Contemporary South Asia*, 15 (3): 325–39.

Fischer, Michael D. and Zeityn, David (1999) *ERA resource guide and sampler CD for teachers and students*. Canterbury, UK: CSAC, University of Kent at Canterbury.

Goody, Jack and Watt, Ian (1963/1968) 'The consequences of literacy'. In Jack Goody (ed.), *Literacy in Traditional Societies*. Cambridge: Cambridge University Press.

Hall, Edward T. (1976) *Beyond Culture*. New York, NY: Basic Books.

Happel, Ruth (2005) 'The importance of place – GPS and photography'. Available at https://web.archive.org/web/20051231125826/http://www.microsoft.com/windowsxp/using/digitalphotography/prophoto/gps.mspx (accessed 16 August 2016).

Kuznar, Lawrence A. and Sedlmeyer, Robert (2005). Collective Violence in Darfur: An Agent-based Model of Pastoral Nomad/Sedentary Peasant Interaction. *Mathematical Anthropology and Cultural Theory* 1(4).   HYPERLINK "http://escholarship.org/uc/item/67x4t8ts accessed 16/8/2016.

Larson, R. and Csikszentmihalyi, M. (1983) 'The experience sampling method', *New Directions for Methodology of Social and Behavioral Science*, 15: 41–56.

Leiner, Barry M., Cerf, Vinton G., Clark, David D., Kahn, Robert E., Kleinrock, Leonard, Lynch, Daniel C., Postel, Jon, Roberts, Lawrence G. and Wolff, Stephen, (2012 [2003]) 'A brief history of the Internet'. In *Histories of the Internet*. The Internet Society. Available at http://www.internetsociety.org/sites/default/files/Brief_History_of_the_Internet.pdf (accessed 16 August 2016).

Macfarlane, Alan (1987) 'The Cambridge experimental videodisc project'. *Bulletin of Information on Computing and Anthropology, Kent University,* Issue 5.

Madakam, Somayya, Ramaswamy, R. and Tripathi, Siddharth (2015) 'Internet of Things (IoT): a literature review', *Journal of Computer and Communications*, 2015(3): 164–73. Available at http://dx.doi.org/10.4236/jcc.2015.35021 (accessed 28 August 2015).

McCarthy, John (1979/1996) 'The history of Lisp'. Available at http://www-formal.stanford.edu/jmc/history/lisp/lisp.html (accessed 8 August 2007).

Moody, James and White, Douglas (2003) 'Structural cohesion and embeddness: a hierarchical concept of social groups', *American Sociological Review*, 68 (1): 103–27.

Owens, A. (2005) 'Semantic storage: overview and assessment'. Technical Report IRP Report 2005, Electronics and Computer Science, University of Southampton. Available at http://eprints.ecs.soton.ac.uk/11985/ (accessed 10 August 2007).

2014   Read, Dwight, Fischer, Michael and Lehman, F. K. (Chit Hlaing). The Cultural Grounding of Kinship: A Paradigm Shift. *L'Homme* n. 210, pp. 63-90.

Ruby, Jay (2005) *Oak Park Stories* (CD-ROM series). Watertown, MA: Documentary Educational Resources.

Sprenger, Polly (1999) 'Sun on Privacy: "Get Over It"'. *Wired*, 26 January. Available at http://archive.wired.com/politics/law/news/1999/01/17538 (accessed 15 November 2015).

Toffler, Alan (1970). *Future Shock*. New York: Random House.

Webster, Frank (1995) *Theories of the Information Society*. London: Routledge.

White, Douglas R. and Johansen, Ulla C. (2004). *Network Analysis and Ethnographic Problems: Process Models of a Turkish Nomad Clan*, Oxford and Lanham: Lexington Books,

Zeitlyn, David (2001) *Reading in the Modern World. Writing and the Virtual World: Anthropological Perspectives on Computers and the Internet*. Canterbury: CSAC Monographs. Available at http://csac.anthropology.ac.uk/CSACMonog/RRRweb./ (accessed 16 August 2016).

Zeitlyn, David and Fischer, Michael (1999) 'Africa divination: Mambila and others'. In M. Fischer and D. Zeitlyn (eds.), *Experience Rich Anthropology*. Canterbury: CSAC Monographs. Available at http://era.anthropology.ac.uk/Era_Resources/Era/Divination/index.html (accessed 16 August 2016).

FURTHER READING

'A brief history of the Internet: by those who made the history, including Barry M. Leiner, Vinton G. Cerf, David D. Clark, Robert E. Kahn, Leonard Kleinrock, Daniel C. Lynch, Jon Postel, Lawrence G. Roberts, Stephen Wolff. In Histories of the Internet, The Internet Society. http://www.internetsociety.org/sites/default/files/Brief_History_of_the_Internet.pdf (accessed 16 August 2016).

Leiner et. al. present a brief history organised around four aspects: technological evolution; operations and management; social aspect; commercialisation aspect.

---

[i]

http://kinsources.net (accessed 16 August 2016)

ii   http://socscicompute.ss.uci.edu (accessed 16 August 2016)

iii   http://hraf.yale.edu (accessed 16 August 2016)

iv   For examples see http://www.webcam-index.com/ and http://www.earthcam.com/ (accessed 16 August 2016).

v   https://www.nike.com/gb/en_gb/p/activity accessed 16 August 2016