eprints@whiterose.ac.uk
https://eprints.whiterose.ac.uk/

The title of the article:

Co-observing the Weather, Co-predicting the Climate: Human Factors in Building Infrastructures for Crowdsourced Data

Authors names:

Yu-Wei Lin, Jo Bates, Paula Goodale

Abstract

This paper investigates the embodied performance of 'doing citizen science'. It examines how 'citizen scientists' produce scientific data using the resources available to them, and how their socio-technical practices and emotions impact the construction of a crowdsourced data infrastructure. We found that conducting citizen science is highly emotional and experiential, but these individual experiences and feelings tend to get lost or become invisible when user-contributed data are aggregated and integrated into a big data infrastructure. While new meanings can be extracted from big data sets, the loss of individual emotional and practical elements denotes the loss of data provenance and the marginalisation of individual efforts, motivations, and local politics which might lead to disengaged participants and unsustainable communities of citizen scientists. The challenges of constructing a data infrastructure for crowdsourced data therefore lie in the management of both technical and social issues which are local as well as global.

Keywords: crowdsourcing, big data infrastructure, citizen science.

Introduction – All weather is local.

In June 2011, the Met Office in the UK launched a crowd-sourcing weather observation website[1] (WOW), in partnership with the Royal Meteorological Society and supported by the Department of Education (Met Office, 2011). Branded as a weather website "for everyone", the WOW project aims to crowdsource weather data from private observers in order to build up a record of weather observations for sites across the UK. The intention of the Met Office, as expressed in a press release, was to "encourage further growth in the UK's amateur weather observing community… help educate children about the weather and…become the UK's largest source of weather observations." (ibid.)

Parallel to this investment in engaging the public in weather observation, the Met Office Hadley Centre for Climate Prediction and Research has also worked with the Zooniverse platform, branded as a collection of "the Internet's largest, most popular and most successful citizen science projects"[2], to initiate the Old Weather (OW) project, which aims to engage the public in the generation of data for climatological science. 'Citizen scientists' are recruited to help recover weather observations made by the crews of historic ships by transcribing digitised versions of ships' log books. These transcriptions contribute to climate model projections and will improve scientific knowledge of past environmental conditions.

These two flagship platforms for crowdsourcing data for atmospheric sciences have attracted much attention, particularly in relation to their technically excellent web-based platforms which enable data collection, and their close connection with the Met Office and other scientific institutions. Undoubtedly, the functionality and interface of the technical systems affects the engagement of potential contributors and/or citizens scientists. However, such a technologically deterministic perspective overlooks how citizen scientists operate and why they participate. Without empirical evidence of how the public, who are the target users of these platforms, perceive the call for their involvement in 'citizen science', and how they engage in these projects and interact with one another and with other stakeholders, it is difficult to develop robust strategies for building an infrastructure for crowdsourced weather data. In turn, this has implications for innovation, knowledge production, and public engagement in science.

This paper addresses these questions from a practice-based perspective by exploring the *glocalised* practices of citizen scientists and the relationship between amateurs and professional scientific experts. Through investigating the experiences and socio-technical practices of amateurs and citizen scientists, we aim to understand the dynamics in the process of building a *glocalised* big weather data infrastructure through connecting various individuals, communities, and organisations through a mixture of bottom-up, organic, modular methods and (semi-) formal institutional management practices. Designed to engage 'everyone', tensions and asymmetries are argued to be found in the construction of these infrastructures for crowdsourcing data. Through investigating the involvement of citizens in scientific research, we also explore the emotional aspect of *doing* citizen science. Challenging the common binary dualisms of the rational and emotional, body and mind, our examination of the experiences of citizen scientists will show that emotions play a major role in motivations. This also advances research on the relationship between amateurs and experts in knowledge production, and on the construction of identities of citizen scientists.

Knowledge Infrastructures

---

1   http://wow.metoffice.gov.uk/
2   https://www.zooniverse.org/about

Various parties (institutions, individuals, communities, organizations), etiquettes, rituals and practices, laws and regulations, facilities and tools are involved in crowdsourcing data. The concept of an 'infrastructure' that contains people, regulations and norms, and artefacts (Star 1999) helps to frame the subject under study as something beyond a technical entity. Several conceptual frameworks proposed in existing STS literature can be adopted to understand the socio-technical dynamics of an infrastructure. For example, it can be epitomised as a unique epistemic culture (Knorr-Cetina, 1999), a community of practices (Lave & Wenger, 1991), a social world where heterogeneous actors and artefacts reside and which has its own hierarchies (flat or tiered), codes, norms, traditions, shared interests, and common practices (Strauss, 1978; Clarke, 1991).

Edwards (2010) provides an infrastructural perspective to understand the development of a global weather and climate knowledge infrastructure. A knowledge infrastructure to Edwards (2010) is a Large Technical System (LTS) where a network of individuals, organizations, artefacts, and institutions are brought together to generate, share, and maintain specific knowledge about the human and natural worlds. This definition of knowledge infrastructures, taking a collection of individuals, organizations, routines, shared norms, and practices into account, echoes Star and Ruhleder (1996), Bowker and Star (1998, 1999), and Star and Bowker's (2010) theories that emphasise the socially constructed aspect of information and communication technologies (ICTs). According to them, infrastructures usually have three components: the artefacts or devices used to communicate or convey information; the activities or practices in which people engage to communicate or share information; and the social arrangements or organizational forms that develop around those devices and practices. These conceptualisations are based on classical STS methodologies and analytical frameworks that call for de-construction and contextualisation of the development and adoption of ICT infrastructures (MacKenzie and Wajcman 1999; Rip and Kemp 1998). They deliver the same message that has been summarised in Edwards et al. (2013), "Transformative infrastructures cannot be merely technical; they must engage fundamental changes in our social institutions, practices, norms and beliefs as well" (p.13).

This paper follows this line of argument by looking into the practices, organisation and manipulation of technical artefacts, and social arrangements within the citizen scientist communities of atmospheric science. These socio-material practices, digital artefacts, and associated norms and rules will be placed in cultural and social-technical contexts where infrastructures like WOW and OW are being developed, organized and governed. But, more importantly, looking at volunteer contributors' practices allows us to uncover those invisible, forgotten, taken-for-granted or hidden figures and issues involved in the construction of an infrastructure for crowdsourced data. This line of investigation is guided by the framework that Star and Strauss (1999) propose in relation to analysing the 'invisible work' of an infrastructure, especially when the infrastructure comprises many sub-systems, each of which is equally complex and within which many practices are made visible and/or invisible. Understanding these visible and invisible practices and processes therefore politicises the development of an infrastructure, and will inform future development of not only the infrastructures themselves (e.g., to improve the engagement with contributor communities, to facilitate easier contributions via better human-computer interfaces), but also of related social theory.

Methodology

The WOW and OW projects were used to frame and scope our study, informing both the collection of empirical data and the sampling of interviewees. Both projects offer a space that enables amateurs (loosely defined communities and/or individuals) to contribute data for atmospheric sciences. The selection of these two citizen science infrastructures was not random. Whilst, WOW is similar to other infrastructures for amateur weather observers such as Weather Underground[3] or the Climatological Observers Link (COL), focusing on the UK-based WOW project and the OW

---

3   http://www.wunderground.com/weatherstation/overview.asp

project, allowed us to examine the local practices and experiences of UK-based amateurs and citizen scientists.

It is also timely to study the WOW and OW projects as the technical systems and the contributor communities engaged in them are still at an infant development stage. As Bowker and Star (1999) note, "Good, usable systems disappear almost by definition. The easier they are to use, the harder they are to see. As well, most of the time, the bigger they are, the harder they are to see.... Infrastructures are never transparent for everyone, and their workability as they scale up becomes increasingly complex" (p. 34). Before the projects get too massive and too difficult to grasp, we aimed to get in early to capture and document as many layers of socio-technical arrangements as possible.

A variety of data have been collected for the purposes of this research, including four in-depth interviews carried out during April-August 2014. Two interviews were conducted with private weather station owners who were potential contributors to WOW, and two were conducted with OW contributors. In the interviews, informants were asked their motivations for collecting or transcribing weather data, the challenges faced, and the enjoyment and frustrations they felt during the processes of, for example, setting up instruments and transcribing data. These interviews were conducted as a part of the *Secret Life of a Weather Datum* project funded by the Arts and Humanities Research Council (UK) during 2014-15. As part of this project, professionals who lead on the WOW and OW projects were also interviewed, and these interviews were used to provide context for the research presented in this article. This wider project aimed to explore the values and practices associated with different projects, organisations and communities on the journey of weather data from initial data production, through quality control and data processing, on into re-use in climate science and financial markets (Bates et al., 2015). The methodology employed, following the spaces, the actors and the evolution of data as a journey, has enabled us to identify and explore the value-making and value-changing processes, and dynamics of components, actors, rules, and relations in the infrastructure. These data were enriched by further data collection including online ethnographic observations on the OW project forum and the WOW mailing list, participatory observations of Maker events, short informal interviews with participants involved in Maker communities, and desk research of documentary evidence relevant to these cases (for example, relevant blogs and press releases). As demonstrated below, these conversations and observations revealed the emotions and bodily performance embedded in the data collection practices, and allowed us to picture the assemblages of a range of actors and objects. The rich narratives collected through the interviews and observations have illustrated different socio-cultural values and practices that shape data production, processing, distribution and re-use on its journey through the infrastructure. The organic yet systematic method of "following a weather datum" (ibid.) exploits the materiality of data, a property Bowker (1994) and Edwards (2010) suggest we should focus on when investigating "infrastructural inversion".

Amateur weather observation and the Weather Observation Website (WOW)

The goal of the WOW project is to engage weather enthusiasts, school students studying weather and climate, and other actors to create an active global online weather community. The kind of data WOW accepts covers a wide range of forms and formats, including ad-hoc information such as notes like 'it is snowing here', or an uploaded photograph of the weather one has observed, or the readings routinely collected from manned or automatic weather stations. It also displays other social media content such as Twitter snow reports tweeted using #uksnow. Website visitors can explore the British weather, looking at how it varies from place to place and moves across the country. A forum has also been established to enable WOW users to communicate with one another, share hints and tips, and to enable the Met Office to provide help and assistance as required[4].

As of 4th April 2013, the MetOffice announced that since launching in June 2011, the website had

---

4   https://groups.google.com/forum/#!forum/met-office-wow

"received more than 100 million weather observations from weather enthusiasts all over the world" (Met Office, 2013). These observations are currently used by the Met Office to provide hyper-local information to meteorologists and UK citizens during extreme weather events, and research is currently being undertaken to explore how the amateur WOW observations might be used for weather forecasting purposes (Bell et al., 2014).

WOW is constantly being improved. For example, it has been updated to make it easier to input observations and photos. The Met Office also has plans to better correlate reporting of weather impacts with associated photos, integrate the Met Office's 5000 weather station site observations into WOW, investigate options for collection and visualization of energy and temperature output data from solar panel systems globally, and improve photo display and search functionality. Users will also be able to submit their observations and photos by mobile phone.

It has been claimed that there was "zero up front infrastructure costs involved, and the platform scales automatically to meet the variable demand from the UK and internationally" (ibid.). This statement on the one hand highlights the easiness and low cost of initiating a crowdsourcing platform, yet on the other hand downplays other factors involved in the development, implementation and maintenance of a socio-technical infrastructure.

Amateur weather observation practices

Many people have weather stations these days (Eden, 2009; Burt, 2012). Commercially available weather stations such as the Davis Vantage are easily acquirable in outdoor or electronics shops on the high street. The Davis consists of fairly standard instruments. It has an electrical resistance thermometer and other standard sensors, a rain gauge on the outside of the station, and some observers also have anemometer to measure wind speed on the roof of their house. The Davis is connected to the Internet, and uploads observation data from the weather station every five minutes (or a different interval configured by the user) to an online data storage platform, which can be downloaded every week or so by the user. Users resultantly have five minute records of a range of variables such as temperature, wind, rainfall, air pressure, humidity, solar radiation etc.

Private weather station owners often have a deep interest in weather observation. As one informant told us,

> "Lots of people have weather stations. It's just a natural thing that if you're interested in something you want to get practically involved, and it's a practical way of getting involved in meteorology and actually measuring the temperature, or measuring how much rain fall. So it makes you understand, it forces you to observe what's happening outside a bit more. And that in turn makes you wonder about the processes and makes you want to read more. So one thing leads to another really. But I like to do things as well as just read about them. So it's really from the practical thing, inclination to really want to immerse yourself in the subject and try and understand more about how things work." [AWS01-1]

In this quote, we can gather that the informant is a self-motivator who enjoys observing and recording weather data.

Bodily performance is highlighted in the following quote from the interviewee, when asked if there are any particular challenges in collecting the data and what can go wrong with it:

> "Obviously, you need to have some familiarity with the equipment to set it up in the first place. It helps obviously, that I had the equipment set up in my previous home. It's always easier setting up something the second time because you're more familiar with it. There is a certain amount of cabling involved because although it's a wireless weather station, I didn't go wireless for all the sensors because it would have been even more expensive. So I had to route some cables from the wind vane and anemometer, and the solar and UV sensors down the chimney, down to the ground, and bury them in the back garden, along a wall and so on. But I've done that sort of thing before. Of course the main challenge is actually mounting the equipment, part of it at a high enough height to record the wind."

[AWS01-2]

Here, we can see the importance of developing one's familiarity with and experience of the instruments and the local environment in order to gather better data. The joy of observing weather goes side by side with the slightly laborious bodily performance of installation and calibration of the equipment.

What does a weather station owner do on a regular basis? It is important to keep a regular and consistent "routine":

> "I don't do as much as I would like to, but I have done. I check the barometer every now and then, at least once a month. And the thermometer I haven't checked for a while, but I actually need to really get hold of a calibration thermometer. The one I've got is pre-calibrated, but that's when I bought it in 2009 and that should really be done once a year. There's a national standard thermometer. I can borrow one, or get hold of one, and then actually just recalibrate really. But in an ideal situation you are meant to recalibrate these instruments every so often, every couple of years I'd say." [AWS01-3]

The opening of this statement is interesting. The respondent seems to know what he *should* do to keep a continuous record or to meet professional standards (e.g., calibrating the instruments), but due to other limitations, he was not able to do so. This on the one hand suggests amateurs' understanding of professional codes of conduct, and on the other hand highlights differences between amateurs and professionals. Whilst the Met Office has to commit to providing accurate and timely weather information, amateurs may have more flexibility, be recording the weather conditions 'just for fun', and feel less obligation to meet professional standards.

The respondent did, however, try to conform to best practices to produce good quality data:

> "You're meant to really calibrate your sensors every now and then because even though it's automatic it's all very easy to leave it just running and assume that the data you're getting are entirely accurate. But of course the data you're collecting are only as good as the instruments that are recording them, which can sort of malfunction or they can show some slow drift in time that might not easily be detectable. In other words they might not be recording entirely accurate data, or they could stop recording if there's some glitch or something. So you need to keep an eye on the data, I'd say on a weekly basis. So that's why the website's useful to keep checking. Occasionally the Internet connection gets lost and then you find it's not archiving the data. But what happens is there's a back up on the weather station, so actually, usually it still is and then you just have to unplug and plug it in a certain way, and take the batteries out and put it all back in. It's a bit of a pain, but it's something that you just have to do occasionally. But it's a pretty good system." [AWS01-5]

In this quote, one learns some ad-hoc local arrangements the private weather station owner developed in order to accommodate local problems or factors. These socio-technical arrangements symbolise "bricolage" (Johri, 2011); one has to make do and adjust to the local conditions faced at that particular moment. They also demonstrate the importance of vernacular and tacit knowledge which is not written in scientific textbooks.

Some of these weather station owners keep the data for their own records, and others share them by uploading onto websites such as WOW, Climatological Observers Link[5] and Weather Underground[6]. Data from thousands of privately owned weather stations are integrated in these various platforms.

The informant expressed excitement about the prospect of using crowd-sourced data to co-produce weather forecasts, and the wider implications of sharing data

> "I'm perfectly happy with having these websites which anybody can access and give a forecast (which I believe, I'm not entirely certain, but I think it's) based partly on my data.

---

5   https://www.colweather.org.uk/index.php
6   http://www.wunderground.com/

There's no point in spending a lot of money on equipment for something I'm passionate about and interested in if it's not in some way benefiting other people, well even from an education point of view. Even you know, the data are not of professional standard, but the station is a semi-professional station so the data can still be used in some research and teaching context, from that point of view. So I mean if it helps Weather Underground with their forecast in a small way, then I'm absolutely fine with that. I think it's great because it's a wider use of the data. So rather than just me using it or my students using it then anyone can log onto the site and use it." [AWS01-4]

This response demonstrates that in some cases, whilst data are being collected because of weather station owners' passion for weather observation, altruistic opportunities for data sharing emerge through time as institutional support evolves and communities of practice grow. Altruism is not essential to the identity of citizen scientists and amateurs, but a quality that is cultivated through the social and technical assemblages they are embedded within. The response also highlights some of the ways in which amateur and professional data and equipment may differ, and points to additional educational and cultural values these amateur-generated data possess. Involving the public in weather observation may encourage citizen scientific culture and improve public understanding of atmospheric sciences. The data can be shared, as long as other socio-technical arrangements, such as web platforms and time, are available.

Whilst the above respondent generated his own weather observation data using a ready-made Davis weather station, some technology enthusiasts build their own weather stations using microcomputers such as the Raspberry Pi. Some participants of Open Source Maker communities such as Raspberry Pi groups, local hackerspaces and FabLabs, and even Linux User Groups (LUGs) have developed an interest in making home-made weather stations. The already diverse and hybrid Open Source Maker communities (Lin, 2005) are further hybridized by such an interplay between citizen science and Open Source making.

An infrastructure that includes the owners of these home-made weather stations and the data they produce, undoubtedly faces challenges of managing, standardising, and integrating different epistemic cultures, especially when amateurs meet experts. We can sense the challenges from the narratives below when the informant discusses their passion for Raspberry Pi technologies. The questions here are: are these different interests (e.g., in the gadget Raspberry Pi or in weather observation) juxtaposed on an equal ground, or is there a hierarchy in terms of preferences amongst them? Do these practitioners consider themselves as 'citizen scientists' or ' Raspberry Pi hobbyists'? In light of the in-depth interview with one Raspberry Pi weather station maker, and informal conversations with participants at other Raspberry Pi makers' events, learning to configure a Pi usually takes priority over weather observation, which is often a secondary interest.

Many of the Raspberry Pi weather station owners are more interested in the low-cost configurable, programmable open-source technological components. Weather stations are one of the classic projects that Raspberry Pi owners build, and various step-by-step construction guidelines can be found in online instructions, technology magazines and books. Building or owning a Raspberry Pi weather station therefore may not necessarily mean that one is interested in weather observation (because if they are interested in weather observation, they may easily get a Davis Vantage, or similar weather station, from the shops). Often, an interest in open source software and hardware co-exist or perhaps outweigh these observers' interest in weather observation. For example, asked what came first - the interest in the weather or the Pi, a respondent who has built not only a AirPi weather station but also done other Pi projects firmly said,

"I was sent a link to the AirPi project essentially and I thought this is very me because it combines several of my previous interests in the form of the electronics, the Raspberry Pi, the weather, programming, er, things I'd done during my degree course. And I thought this seems like a very nice way to try meshing knowledge in a new way." [AWS02-1]

Members in such Maker and Hacker communities often express that they build or collect things 'just for fun' (e.g., Torvalds 2002). This emotional expression requires a deeper understanding – fun for whom? Why is it fun? Why would or wouldn't a Raspberry Pi weather station owner contribute

the data to WOW? Is it because it is less fun? Where does the fun part end – if at all?

These are interesting questions with regard to motivations, but they also relate to the materiality and affordances of the Raspberry Pi. Asked what he enjoyed about having a Raspberry Pi, a weather station, and the resultant data, the respondent said,

> "It's kind of my version of art. People paint as creative expression, my creative expression is a bit more logical in terms of programming. I always quite enjoyed Lego as a kid and, specifically what I enjoy is the constrained solutions - if you're trying to do something and you have these resources how can you best do what you're trying to do? And so building the weather station is kind of a subset of that but it's why I get into a lot of programming of electronics. I got this neat idea how can I do it with what I already have or getting the least amount of stuff possible off eBay and things like that. And so the Raspberry Pi weather station is just another version of that." [AWS02-2]

Richard Stallman, the founder of the Free Software Foundation, became a free software advocate and practitioner because he wanted to fix a paper jam, a very personal and local problem (Williams 2002). Similar to Stallman's paper jam problem, and the findings from numerous free/open source software studies (e.g., Ghosh et al., 2002, Lin, 2005, Freeman, 2007), the motivation for turning a Raspberry Pi into a weather station In this case can be attributed to solving an existing problem at hand:

> "I had the barometer because I was getting quite tired of the let's go check BBC weather. For short term predictions, I can generally get a good idea of what's happening off the barometer." [AWS-2-2]

Our respondent had no plan for sharing his data with anyone, uploading them anywhere, or doing any analysis of them. He said that he had managed to have the weather station recording since January 2014, so six or seven months data existed at this point.

> "I don't have any definite plans because for me that weather station is hobby territory not must absolutely do it work territory. And so I'm just sort of enjoying the graphs and the nice little thing in the corner of my screen on my desktop PC which shows the latest readings there as well. I'm just sort of enjoying those things and be able to check if it's been raining and what does the rainfall look like?" [AWS02-3]

This problem-solving mindset and behaviour also leads the respondent to disregard himself as a 'citizen scientist'. To him, he was only interested in trying out and adding different sensors onto the Raspberry Pi for "a good learning experience". He recounted:

> "For me I wouldn't class too much of what I do as citizen science. I mean the Raspberry Pi stuff that I write about you could count as 'educational science'. I would class something as potentially citizen science if someone was applying his professional knowledge to doing it. I know I am not." [AWS02-4]

Whilst the respondent, who is an open source software developer and advocate, didn't currently share his weather observation data via a platform such as WOW, drawing on his open source experiences he did recognise that he would get some benefit from doing so:

> "The motivation for sharing the data I suppose would just be a cross between… something along the lines of I've got it I might as well share… crossed with, er, trite, but sharing is caring sort of thing… You do get a little bit of a… not jolt, but boost, or you get a little visceral pleasure from sharing and helping other people out and it would come under that." [AWS02-5]

When questioned why he did not share the data he collected, the respondent explained that whilst he shared his software code, he was concerned that the quality of his data was not good enough for sharing. Further, whilst he was open to considering sharing data for some weather variables he thought were more accurate, he didn't feel it was a priority for him at the present time:

> "I have been considering doing that for the things which I know wouldn't be affected by the

sunlight so that's particularly with the pressure and for the rainfall but also means I do have to write then the software model to do that. And it's not hugely complex I just haven't got into the right frame of mind where I'll sit down and write this bit of software today. So I haven't done it but in the future I suppose I would be interested in doing that because it does seem interesting" [AWS02-5]

The challenge of 'time' again is flagged up here. If the informant doesn't have time, it is difficult to make commitments and provide consistency in data collection or tool improvement. The practitioners may have interests and motivations, but 'time' is a critical factor that affects their engagement.

This view is quite common amongst those who are engaged in this wider hackers' community, loosely structured by members who share a repertoire of open source practices (Lin, 2005). Even if the Pi weather station owners have demonstrated that they can collect data, and they believe in open source philosophy, they don't necessarily prioritise sharing the data they have been collecting. Their motivation for collecting data is not necessarily because of concerns about meteorology or climate change, but something 'tokenized', something linked with practicality, passion, and emotions, rather than altruistic 'gifting' to the wider community. Phrases such as "just in case one day I need it", "just for fun", "just because I want to" and "just because I can" were heard often in informal conversations at Maker events.

Climate data rescue and the Old Weather project

> "It's the weather, it's the history, and it's the forum I think for me are the three key important things that have sort of kept me interested in it really." [OW1-20]

The Old Weather project was initiated to help climate scientists use weather data from historic ship log books to study climate patterns from the past. Before satellites, weather data transmitters, and computer databases, weather conditions at sea were dutifully documented by sailors by hand in the log books of ships. For years, climate scientists have been keen on using these historical records to establish baseline climate data. However, much of these data exist only in hand-written documents stored in archives and are inaccessible to most people.

Dr. Philip Brohan, a climate scientist at the Met Office Hadley Centre since 2002, has been leading the Old Weather project that crowdsources efforts to transcribe scanned copies of log book pages, some more than 150 years old, and make them available to climate scientists worldwide (Brohan et al., 2009). Project scientists integrate the transcribed data produced by Old Weather volunteers into existing large-scale data sets, such as the International Comprehensive Ocean Atmosphere Data Set, which are used by researchers around the world. Begun in 2010, in its first two years the Old Weather project involved more than 16,000 volunteers in transcribing 1.6 million weather observations from British Royal Navy log books.

As well as weather observations, the log books also contain information on maritime history, scientific explorations, military operations, and dramatic rescues and shipwrecks at sea. While the data extracted from these records will be useful to climate scientists, these documents are also a wealth of information for historians, genealogists, people who wish to find out their family histories, or anyone interested in exploring the diplomatic, scientific, technological and military aspects of the voyages, and the experiences and accomplishments of seafaring people.

Because of its intersection with historians and maritime enthusiasts, the Old Weather project engages a diverse group of volunteers (or 'citizen scientists') (Eveleigh et al., 2014), quite different from the amateur weather observers or the Raspberry Pi Makers community. One informant who has been involved in the project for nearly four years told us that she learned about the project on BBC Radio 4. She was rather taken by the idea of contributing to climate science to address climate change. The other informant, an administrator in an Environmental Science department in a UK university who has also been involved in the project for more than three years, said she was moved partly by her curiosity about her colleagues' work, and partly taken by her concern for the planet. It was this "wider picture" that kept her hooked for so long:

> "Feeling that that is a worthwhile thing to do, and it's contributing to a scientific project that I think is important. And then I think I got interested in the wider picture as it were, of life on board the ships, and the whole thing of the naval history mostly of the First World War, about which I knew nothing. So it kind of spread itself out into all the other topics as well."
> [OW1-1]

A social conscience, some background knowledge in weather observation (some even run their own weather stations), and interest in history are widely shared amongst the participants. Each of these three elements are linked with motivations and are highly emotive. Those emotions are clearly demonstrated in the accounts the informants provided, especially with regard to their interaction with the historical materials and with fellow participants.

The historical data, for example, contain certain narratives that move people. Volunteers experienced emotions by reading the log books, and feel attracted to the historical materials they view online. Reading and transcribing these historical materials also give volunteers a sense of connection to the lives of people that lived many years ago. As one participant vividly described:

> "I don't know how but it does feed into one's imagination, and a broader sense of sympathy. On one of the ships I was on, it was coming back from Africa after the First World War had ended. And the number on the sick list kept going up, and of course it was the influenza epidemic. And I remember realising that I was really quite anxious about this ship and this crew. I was thinking this is silly, you know, this is all a very long time ago, whatever's happened's happened. But I realised I was really getting quite anxious about my crew, and you know, hoping that they were all going to, you know having come through the war that they were actually going to come through the flu epidemic." [OW1-2]

Transcribing historical data therefore is not a mechanistic act. It is embodied, emotional, personal, and connected with one's interests and existing tacit knowledge of histories and geographies. Telling the interviewer what she chose to transcribe, an informant said:

> "The Royal Navy ones after a bit I got that there were certain parts of the world I quite liked, and other parts of the world I was less keen on. So if I'd finished one ship and was looking for a new one I quite often thought I'd like another one that is for example, in East Africa because I'd done one or two there, and I'd got to know the names of places, and all that kind of stuff." [OW1-3]

The Old Weather project, as also seen in the case of amateur weather observers, confirms again that 'citizen science' involves highly embodied and emotive activities. When volunteers were asked to work on newly digitised North American ship logs introduced in 2012 after the success of transcribing Royal Navy Ships' logbooks from the period around the First World War, there was some initial resistance. Problems occurred during this period because these emotional and embodied dimensions weren't fully recognised. Some volunteers deliberately avoided transcribing these new materials. This is because many of the volunteers had little knowledge about the American ships and histories, and it appeared to be intellectually as well as emotionally difficult for them.

> "It was really quite hard work because the American logs were very different to the Royal Navy ones. The interface was also changing. The initial interface was really quite experimental, and it was just very hard going." [OW1-4]

This change in the source of materials being transcribed – the result of a celebrated collaboration between The National Archives (UK) and the National Oceanic and Atmospheric Administration (USA) – had a dramatic impact on community dynamics and practices:

> "With the American boats being different, the databases working very poorly, the frustration of how bad it was at various things... The poor moderators had to keep everybody happy because at that point [name of former participant] had gone, we'd had some fun, it was all looking like a disaster, we were in the unfamiliar zone, and it would have been very easy then for everybody to go. But somehow we got ourselves through that. Then it was a case of everybody trying to be as jolly as they could, keep the things going, lauding the work that

we were doing so far. Picking up interesting things from the American ships to try and make them look as interesting as the Royal Naval ones had been. But I think we were on a knife edge at that particular moment, it was very scary. We did lose a lot of people who decided that actually, the whole thing meant so much to them that to cut and run was probably the only sensible way to deal with it. And there's people like me who actually can't imagine life without it. [OW2-3]

This informant has used a lot of (negative) emotional words in this extract, such as 'frustration', 'un/happy', 'disaster', 'unfamiliar', 'trying', 'scary'. This extract reveals the affect the expanding Old Weather data infrastructure imposed on her and other participants. Another recounted:

"Because there was a big change when the American ships came in, and a lot went, "Oh it's nothing like the Royal Navy books, I don't really understand what's going on here." And this off switch of comfort just said this is not the familiar anymore, this is not what you chose to do, but what you did like doing was the editing, and there's tons of that left. So a lot of people said, "I think I've done my bit for citizen science climate transcriptions, let somebody else have a go and I'll go off and do my editing," which takes a certain amount of experience to do I think." [OW2-4]

Here, we see how the change of the OW infrastructure (the involvement of new institutions, larger databases and a new interface) shapes the community practices, attitudes, behaviours, and dynamics. A loss of the 'familiarity' experienced with the Royal Navy materials and histories, generated uneasiness and discomfort for the participants. While many technologists would consider "the more data the merrier" in a big data era, the data from the field demonstrates that the OW community members had mixed feelings about the addition. Even if the citizen scientists understood the purpose and usefulness of the American ship logs - "[At the phase when] the American logs were chosen specifically to provide weather records for, particularly for the Arctic, and that sort of part where they didn't have many records. So they looked for where they were lacking, and found ships that would provide that, so it's very targeted" - the participants could not help feeling alienated from the new log books from the American ships. The negative emotional response to certain types of data to be added was due to their attachment to certain historical materials, personal knowledge of specific historic periods and regions, confidence of rendering accurate and credible data, and familiarity with original materials. Not being as familiar with the history of North America and the new materials, made it initially more difficult for them to engage, transcribe, and edit the ships' log books. Nonetheless, over time many of the participants adapted to the change, and pushed ahead with the transcription task.

These subtle and often hidden relationships between data and data users are hinted at by Bowker (2005) when he proposes that "raw data" are an "oxymoron". Following this argument, others such as Gitelman (2013) have rejected the presumed objectivity of data, arguing that data afford certain types of knowledge to be produced, rather than innocently discovered. We subscribe to these arguments, and consider the relationship between data (the original inscriptions recorded in the ship log books as well as the value-added data produced through different processes) and citizen scientists' emotional responses and sentimental feelings towards data. As argued earlier, the narratives and textuality of these historical records have driven the volunteers to engage with and rescue the stories of the ships' crews. The value-added data generated by the volunteers of the Old Weather project therefore are not just fact-based scientific weather records, but also other accounts of everyday life and occasions including death. These narratives are not trivial, but impact different lives in a variety of ways.

Asked to assign values to the voluntary work she has been involved in and compare them, one respondent reflected:

"I think the scientific value I would put first, but then definitely the historical information, which is also being recovered, in terms of the other comments in the logs. And I think particularly stuff about people. We fairly regularly get people posting on the forum saying, I am researching my family tree and I know that my grandfather, or my great uncle, or whatever was on this ship, you know is there any record of him? And we're able to point

them, perhaps to the logs or to say, "they're not up yet, but they should be, so check back", this sort of thing. So I think it's helping to recover some history that isn't going to get known about otherwise. And actually, sometimes correcting information, which has been slightly wrong, for example deaths in particular 'cause we start recording all the deaths of anybody. Now the majority of them were already recorded, but sometimes the information we had from the log was actually a bit different in terms of cause of death, or the date, or whatever. And also we've sometimes had recordings of deaths of people who were part of the crew, but weren't actually naval personnel - boys who were sort of local, in Africa particularly, who were taken on board, and they tended not to get recorded. There were a few where it was actually recorded, a death, and so we've made sure that they get kept. So there's a bit of sort of almost recovery of lost history in some ways. Which also feels important to me, and kind of honouring people in a sense. Particularly in the people sense of it that honouring people who you know, perhaps died of this and maybe haven't been recorded at all. We can add a bit of detail perhaps, particularly if they were buried at sea we might be able to actually have the location for example because they did quite often put in the latitude and longitude when they buried somebody at sea." [OW1-5]

Some of the historical value of the OW data, especially interest from external people such as members of the public who had ancestors on the ships or originating from different continents were unexpected by some of the OW participants. However, these observations demonstrate the ways in which these crowdsourced data are not confined to scientific interpretation, but are also open to a wider, more diverse, use and interpretation. These historical data are collated through an editing process, and are shared via the naval-history.net website for anyone to access and read.

The embodiment in *doing* 'citizen science' can also been seen in the hidden, invisible, and often emotional practice of reading and making sense of hand-written historical documents. For example, flagging up the problem of transcribing digitised 'handwritten' historical documents, where the handwriting varies enormously, one informant shared her frustration saying,

> "[The handwriting] can vary a lot even just on one page; you can get half a dozen different handwritings on one page of a log sometimes. I think definitely one of the main frustrations is just trying to decipher what it is, and trying to make sure, particularly with the weather records that you're as accurate as possible because three people have to transcribe each page. … If everything is different then that weather record basically isn't useable, it gets thrown out because it's not accurate enough. You really are wanting to make a big effort to get it as accurate as you can, and hope that everybody else is too." [OW1-6]

The accuracy of the data was emphasised in the quote above. To ensure the data accuracy, the participants have to familiarise themselves with not only the instructions but also the social norms of asking for help on the forums. For example, how to ask and frame a question:

> "Particularly with editing, I usually go through a reasonable amount of the ship and then I start posting questions, sometimes about odd things I haven't been able to either read, or I think I can read it, but I've no idea what it means. Does anyone know what's going on here as I've been unable to find anything?" [OW1-6-1].

Socialisation is a good way of learning and finding solutions to overcome the problem of discerning handwriting. Our forum observations and the interview data suggest that most of the socialisation took place online rather than offline. Zooniverse organises annual conferences for volunteers to meet up, but it was the forum that played an important part in many volunteers' life and was mentioned again and again in the interviews. An Old Weather participant said,

> "It's quite unusual, it is pretty much all online. There's a facility to send personal messages, so some of it isn't an open forum. It's not just you sitting at your computer in isolation transcribing away. It's also actually relating to other people who are doing it, and assisting them, being able to ask for assistance. ... And quite often other people can come up with something. There are one or two people who are absolutely brilliant at tracking down obscure ships, for example. And others who've got a really good eye for odd handwriting. Or just people who happen to know that part of the world, for example, and therefore you

know, are more likely to be able to work out where are we, what is this name, or whatever. So it kind of draws on everybody's skills I think. Sort of a group effort." [OW1-7]

The personal and tacit knowledge has been highlighted in this quote. This echoes what is mentioned earlier about the role of local and tacit knowledge of an amateur weather observer. Asked what kept her motivated overall in what she did with the Old Weather Project, another informant said

"I think the sense of contributing to something that I care about, but also definitely the forum. The forum is massively important. It's an extremely useful source of information and assistance. But it's also a real community. I was just looking at it before our chat, having a look to see what had happened since yesterday, and in the chat thread someone has just announced the birth of his first child, for example, one of the transcribers [laughs]. And we have that quite a bit. People are telling each other about important things in their lives, or that they're going off on holiday so they won't be around for a bit, but they'll put some photographs up when they come back, and this kind of thing. So it's got a real kind of community sense, as well as being a very useful source of can anybody read this writing, does anybody know what's happening here." [OW1-8]

The online social space was described as "a very friendly place" with "a support element to it [plus] a lot of personal interaction as well as some fun bits" [OW1-12]. One informant who had also tried other citizen science projects on Zooniverse explained why she favoured the 'Old Weather' project:

"There's the opportunity to be more involved; the opportunity to have both the social life and getting the citizen science out of things is there, and that's the mix that I like. Whereas some of the others like the Mars stuff just seemed empty, barren, devoid of personality really, and that does not suit me." [OW2-2]

Crowdsourcing data infrastructure and connected communities of practices

Data can be scaled up, through some form of organization, standardization or institutionalization, to become 'boundary infrastructures' (Bowker & Star, 1999). Extended from the original idea of a "boundary object" (Star & Griesemer, 1989; Clarke & Fujimura, 1992) through which diverse actors are brought together to shape and interact within a large platform or infrastructure, we can conceptualise these crowdsourced data objects as a form of boundary object that connect different individuals and communities as they move through the infrastructure. In this sense, the crowdsourced data infrastructure should be recognised as a "glocalised" socio-technical infrastructure, containing various 'boundary data objects' whose production, processing, distribution and use are embedded in local practices and value systems that resonate with local conditions and limitations.

This modular way of building and connecting communities of practices enacts the 'scalability' and 'extensibility' of a big data infrastructure (boyd, 2014; Kitchin, 2014). However, it's important to acknowledge that when a data infrastructure expands, not only data but also a range of socio-technical elements are assembled. These modularized components include communities, tools, pathways, and methods. In the communities we study here, in which the general public are connected with the professional scientific community, additional challenges are also brought into play in relation to the management of scientific knowledge production:

1. Local, personal, and tacit knowledge

The fact that there were far fewer people transcribing the American ship logs (compared to the number of volunteers working on Royal Navy's ship logs), and that many felt "This is not my cup of tea", emphasised that different citizen science projects are attractive for different types of people. The motivations for getting involved vary from individual to individual. It is very personal and very embodied. Deeper engagement with citizen science requires local knowledge, interests and emotional attachment – something participants can associate with and recognise cultural references or interests.

2. Socialisation

Having a shared place for mutual support or knowledge sharing is another crucial feature in citizen science projects. This may take forms of face-to-face real-life meet-ups (e.g., Zooniverse annual meetings or Makers faires) or on-line forums or mailing lists. Raspberry Pi makers' communities self-organise many online forums to support one another and facilitate cross-boundary learning and problem solving. Members of the OW community tend to favour conversations that take place on the project's online forum, perhaps more so than the WOW mailing list members. Our observations of the OW forum found a lot of light-hearted dialogues illustrating community support and social interaction.

3. Embodiment (the physical, emotional and cognitive activities involved in recording, observing, transcribing and editing)

Weather observation involves more than recording scientific facts. Transcribing and editing historical records also requires more than just reading and typing. In the former, configuring and tinkering devices is a common practice found amongst amateur weather observers. In the latter, OW citizen scientists have engaged with recovering data and stories, empathising with and caring for historical shipping crews, imagining seafarers lives, and guessing old-fashioned handwriting. Understanding some of the hand-written documents was the biggest challenge some OW informants reported. There were times people had to 'improvise': "We're all told that if you really can't read it, guess extravagantly because actually, you probably know better than anybody else what it's likely to be if you've been transcribing for a while" [OW1-9]. We can therefore recognise that crowdsourced data are inscribed with emotions, experiences and bodily performances.

4. Attitudes towards professional standards and data quality

As seen in the narratives provided by the amateur weather observers and the OW participants, the citizen scientists we interviewed were aware that the weather data they produced might not be 100% accurate. However, desires for the quality of data that expert scientists strive for were nonetheless reflected in the volunteers' practices and mind-sets. OW respondents, for example, demonstrated a strong sense of duty to the project – emphasising a desire for completeness and accuracy. Mechanisms (formal and informal) were developed to ensure data quality and standards. For example, to ensure the accuracy of the transcribed data OW volunteers peer-review one another's work, and the amateur observers took time and efforts to calibrate their instruments and data to take local conditions into account. Aware of the importance of good quality data, most of the volunteers had a strong sense of responsibility with regard to the data they were producing.

5. Trust from the professional scientists

The relationship between citizen scientists and the professional expert scientists provides insight into the citizen scientists' attitude towards their roles and responsibilities, and their self-identity as participants on projects such as Old Weather. The volunteers' dedication to completeness and accuracy garnered respect from the climate scientists, who spent time engaging with and building relations with members of the community and answering questions if needed. The interview data suggests a genuine sense of responsibility and delight is generated through interactions with the professional climate scientists.

Given the diversity and heterogeneity within and across these citizen science projects, a crucial question for understanding a big data infrastructure based upon them is how to homogenise and integrate these crowd-sourced data collected and generated in distributed environments into a global big weather and climate data infrastructure. This is not merely a question of 'how to' achieve this technically, but also one of how to tackle the social issue of ensuring that the diverse interests existing in different citizen science projects are harmonized, sustained and maintained within a single infrastructure.

The existing STS literature has addressed the issues regarding homogenizing and standardising boundary objects (see e.g., Star & Griesemer, 1989; Fujimura 1991; Wenger 2000; Lee 2007; Star 2010; Jensen and Kushniruk 2014) but the issues haven't been discussed in the context of distributed data collection and generation. Our study begins to bridge this gap by looking into the construction of infrastructures for crowdsourced data, in a similar effort as seen in the two articles published in the earlier parts of this special issue on knowledge infrastructures: the production of Wikipedia (Wyatt et al., 2016), and grassroots infrastructures (Jalbert ,2016). The aforementioned communities of practices (weather enthusiasts, private weather station owners and citizen scientists), though seemingly unrelated, all share one character, which is a loosely defined (and perhaps also ephemeral) boundary and a flexible membership. Members of these communities have varying interests. The data are inscribed with the contributors' memories of places, lifestyles, interests, values, and communities they reside in. In a big data infrastructure where the data crowdsourced from different origins are aggregated and integrated, these data that are produced by different parties dislocate from the places they came from.

Although data are usually considered as scalable and extensible in a big data infrastructure, our findings suggest that, whilst scalability may be relatively achievable on the technical side, it is more difficult to handle aggregated, augmented, and accumulated human factors on the social side of the infrastructure, especially in relation to people's emotions, memories and attachment to histories, norms, traditions, and social spaces. While the data can be aggregated, the memories and emotions and human factors cannot be accumulated at the same scale, speed, or in the same way. When data are put together, the personal characters of these data are erased. From our investigation into those hidden and invisible practices of citizen scientists involved in the OW project, for example, we found the challenge of dealing with human factors in a scalable big data infrastructure. Participants reported the struggle of maintaining motivations when the materials being transcribed became disconnected from their personal interests and existing knowledge base. Building up a big data infrastructure involves not only aggregating data, but also human factors. These hidden issues can only be identified if we understand the local practices of data generation and collection, how they shape the ecology of the infrastructure, and what the 'matters of concern' are for those invisible workers who take care of infrastructural breakdowns, failures, and repairs (Star, 1999; Star & Strauss, 1999).

Conclusions

While crowdsourcing user-generated and user-contributed content and data has become an accepted method for producing scientific knowledge, it is timely and important to get a better understanding of how crowdsourcing data infrastructures operate. In these kind of large-scale, networked computing infrastructures where data that are generated and collected from different sources are housed, processed, and aggregated, the 'bigness' has been seen in terms of quantity as well as variety (formats, types). Data included in such crowdsourced big data infrastructures come from various sources, and are generated by different users and organizations through different means. All these data collected, collated and generated in different ways for different purposes denote diverse (and sometimes conflicting) agendas and identities, materialised in specific forms that can be converted into different formats that are re-used, re-mixed, aggregated, re-contextualised, and re-purposed. To understand the construction process of a big data infrastructure, we need to understand how these diverse communities, individuals, organisations and institutions function at the local level and the outcomes and consequences when they are connected together.

This paper has looked into the local experiences and practices of amateur and citizen scientists contributing to atmospheric sciences. The respondents in this study include amateur weather observers who create their own digitalised records of the weather, and citizen scientists who contribute to the OW project to restore and recover historical archive materials. We have highlighted the affective and emotional aspects of the practices and bodily performance to tease

out the visible and invisible human factors involved. We have also discussed the challenges of dislocating and depersonalising these crowdsourced data in a big data infrastructure, especially in terms of loss of motivations and sense of identities.

Whilst a scientific data infrastructure often denotes something more stable, standardised, structural, and institutionalised, the involvement of citizen scientists creates a more unstable and uncertain space. How to coordinate and sustain the efforts of these diverse communities and integrate them into a big weather data infrastructure remains a challenge to be overcome.

References

Bates J, Goodale P, Lin Y-W (2015) Data Journeys as an approach for exploring the socio-cultural shaping of (big) data: the case of climate science in the United Kingdom. iConference 2015 Proceedings. URL (Retrieved on 1 Feb 2016): http://hdl.handle.net/2142/73429.

Bell S, Cornford D & Bastin L (2013) The state of automated amateur weather observations. *Weather* 68: 36-41.

Bowker GC & Star SL (1998) Building Information Infrastructures for Social Worlds - The Role of Classifications and Standards. *Community Computing and Support Systems*, Vol 1519 of the series *Lecture Notes in Computer Science*, pp. 231-248.

Bowker GC & Star SL (1999) *Sorting things out: Classification and its consequences*. Cambridge, MA: MIT Press.

Bowker GC (2005) *Memory Practices in the Sciences*. Cambridge, MA: MIT Press

boyd D & Crawford K (2012) Critical questions for big data. *Information, Communication and Society* 15(5): 662-79.

Brohan P, Allan R, Freeman JE, Waple AM, Wheeler D, Wilkinson C & Woodruff S (2009) Marine observations of old weather. *Bulletin of the American Meteorological Society*, 90(2): 219-230.

Burt S (2012) *The Weather Observer's Handbook*. Cambridge: Cambridge University Press.

Clarke AE (1991) Social worlds/arenas theory as organizational theory. In D. Maines (ed.) *Social organization and social processes: essays in honour of Andelm L. Strauss*. NY: Aldine Gruyter.

Clarke AE & Fujimura JH (eds.) (1992) *The Right Tools for The Job: At work in Twentieth-Century Life Sciences*. Princeton, NJ: Princiton University Press.

Eden P (2009) Traditional weather observing in the UK: an historical overview. *Weather* 64: 239–245.

Edwards PN, Jackson SJ, Chalmers MK, Bowker GC, Borgman CL, Ribes D, Burton M, & Calvert S (2013) *Knowledge Infrastructures: Intellectual Frameworks and Research Challenges*. Ann Arbor: Deep Blue. http://hdl.handle.net/2027.42/97552.

Eveleigh AMM, Jennett C, Blandford A, Brohan P, Cox AL (2014) Designing for dabblers and deterring drop-outs in citizen science. CHI'14: Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. pp. 2985-2994. New York, NY: ACM.

Freeman S (2007) The Material and Social Dynamics of Motivation: Contributions to Open Source Language Technology Development. *Science & Technology Studies* 20(2), pp. 55-77.

Fujimura JH (1992) Crafting science: standardized packages, boundary objects, and "translation". in Andrew Pickering (ed.) *Science as Practice and Culture.* Chicago: the University of Chicago Press.

Ghosh RA, Glott R, Krieger B & Robles G (2002) Free/Libre Open Source Software: Survey and

Study - Part 4: Survey of Developers. Deliverable D18: Final Report. International Institute of Infonomics University of Maastricht, The Netherlands.

Gitelman L (ed.) (2013) *"Raw data" is an Oxymoron*. Cambridge, MA: The MIT Press.

Jalbert K (2016) Building knowledge infrastructures for empowerment: A study of grassroots water monitoring networks in the marcellus shale. *Science & Technology Studies* 29(2), pp. 26-43.

Jensen S & Kushniruk A (2014) Boundary objects in clinical simulation and design of eHealth. *Health Informatics Journal*, 22(2), pp. 248-264.

Johri A (2011) Socialmaterial bricolage: The creation of location spanning work practices by global software developers. *Information and Software Technology*.

Kitchin R & Lauriault TP (2014) Towards critical data studies: Charting and unpacking data assemblages and their work. In Eckert, J., Shears, A. and Thatcher, J. (eds) *Geoweb and Big Data*. University of Nebraska Press.

Knorr-Cetina K (1999) *Epistemic cultures: how the sciences make knowledge*. Cambridge, MA: Harvard University Press.

Lave J & Wenger E (1991) *Situated learning: Legitimate Peripheral Participation*. Cambridge: Cambridge University Press.

Lee, CP (2007) Boundary negotiating artifacts: Unbinding the routine of boundary objects and embracing chaos in collaborative work. Computer Supported Cooperative Work, 16, 307- 339.

Lin YW (2005) *Hacking Practices and Software Development: A Social Worlds Analysis of ICT Innovation and the Role of Open Source Software*. PhD Thesis. York, UK: University of York.

MacKenzie D & Wajcman J (eds.) (1999) *The Social Shaping of Technology*. Milton Keynes: Open University Press.

Met Office (2011) WOW - A new weather website for everyone. The Met Office News Release on 11 February 2011. URL (retrieved on 13 October 2014): http://www.metoffice.gov.uk/news/releases/archive/2011/wow-website-to-launch

Met Office (2013) WOW hits 100 million observations. The Met Office News Release on 4 April 2013. URL (retrieved on 13 October 2014) http://www.metoffice.gov.uk/news/releases/archive/2013/WOW-hits-100million

Rip A & Kemp RPM (1998) Technological Change. In S. Rayner, E.L. Malone (eds.), *Human Choice and Climate Change*. Vol. II, Resources and Technology. Columbus, OH, Battelle Press, pp. 327-399.

Star SL (1999) The Ethnography of Infrastructure. *American Behavioral Scientist*, 43(3): 377-391.

Star SL (2010) This is Not a Boundary Object: Reflections on the Origin of a Concept. *Science, Technology & Human Values*, 35(5), pp. 601-617.

Star SL & Bowker GC (2010) How to infrastructure. In L. A. Lievrouw and S. Livingstone (eds). *Handbook of New Media: Social Shaping and Social Consequences of ICTs*. Sage.

Star SL & Griesemer JR (1989) Institutional ecology, translation and boundary objects - amateurs and professionals in Berkeley's Museums of Vertebrate Zoology, 1907-39. *Social Studies of Science*, 19(3): 387-420.

Star SL & Strauss A (1999) Layers of Silence, Arenas of Voice: The Ecology of Visible and Invisible Work. *Computer Supported Cooperative Work*, 8 (1), pp. 9-30.

Star SL & Ruhleder K (1996) Steps toward an Ecology of Infrastructure: Design and Access for Large Information Spaces. *Information Systems Research*, 7(1), 111-134.

Strauss A (1978) A social world perspective. *Studies in Symbolic Interaction* 1: 119-128.

Torvalds L & Diamond D (2001) *Just for fun: The Story of an Accidental Revolutionary*. London:

HarperBusiness.

Wenger E (2000) Communities of Practice and Social Learning Systems. *Organization* 7(2): 225-246.

Williams S (2002) *Free as in Freedom: Richard Stallman's Crusade for Free Software*. Sebastopol, CA: O'Reilly.

Wyatt S, Harris A & Kelly SE (2016) Controversy goes online: Schizophrenia genetics on Wikipedia. Science & Technology Studies 29(1), pp. 13-29.