The Citizen Scientist in the ePolicy Cycle

Johann Höchtl, Judith Schossböck, Thomas J. Lampoltshammer, and Peter Parycek

Abstract This chapter discusses a participation and technology enabled model of the citizen scientist in relation to the policy cycle. With interconnected personal devices collecting a plethora of various data, citizens are capable to serendipitously contribute to crowded knowledge generation. In the governance domain, the trend towards more data-driven models of governance and decision-making has been considerable. Big data contains the methodologies to cope with the wealth of data generated by the citizen scientist and in turn provides the tools and technologies to draw actionable insights from this data, f.i. with predictive technologies that could optimise resources across government sectors. After discussing the changing role of science and the technological and participative enablers and methods of engagement relevant for citizen participation, this contribution discusses the role of the citizen scientist and his or her involvement in the big data enabled governance loop by defining three use cases within the policy cycle. Furthermore, it addresses the challenges that can arise in this context.

Introduction

The term science as well as the nature of conducting science evolved over time. Not always has research revolved around the methodological approach as we know it, and not always has it been driven by the measures of today. In this paper we start by describing the nature of conducting science and how some scientific paradigms changed over time. This is relevant for our analysis of citizen science in relation to the ePolicy cycle, as changes like the focus on the openness paradigm combined with available means for sharing and mass collaboration also changed how citizens can participate in the research process.

The section Citizen Science focuses on how openness in the research process combined with means for mass collaboration can empower citizens to enrich the

J. Höchtl • J. Schossböck • T.J. Lampoltshammer (✉) • P. Parycek
Department for E-Governance and Administration, Danube University Krems, Dr.-Karl-Dorrek-Str. 30, 3500 Krems, Austria
e-mail: Johann.Hoechtl@donau-uni.ac.at; thomas.lampoltshammer@donau-uni.ac.at

A. Ojo, J. Millard (eds.), *Government 3.0 – Next Generation Government Technology Infrastructure and Services*, Public Administration and Information Technology 32, DOI 10.1007/978-3-319-63743-3_3

research arena. After describing these changes, we take a more detailed look at the possible ways to engage citizens in this process. The section Enablers and Methods of Citizen Engagement summarises some recent modes of citizen participation and engagement, mostly in relation to ICTs and digitalization, and the participatory and technological aspects of citizen engagement. We also briefly address the opposite of those enablers in the form of hurdles to citizen science.

In the Big Data Enabled Policy Cycle we present the policy cycle as a theoretical vehicle to structure public policy making in light of technological advance. Use Cases for the Citizen Scientist in the Policy Cycle ties together intrinsic motivation and external enablers in respect to the policy cycle. In Challenges, Issues and Future Implications we discuss existing impediments to unleash the potential of Citizen Science in policy making, ethical and cultural considerations as well as potential implications of future research.

By combining insights from different disciplinary fields, we hope to point towards the chance of engaging citizens on various stages of the policy cycle, in particular with view to an increased culture of sharing and related possibilities for evidenced-based and participatory policy making.

Changing Paradigms in Science

For the most part of history, science was not meant for everyone. In former times, many people lacked the basic foundations of what was perceived to be a pre-requisite for scientific work, namely mathematics, jurisprudence, medicine, theology, and philosophy. The *lingua scientia* was dependent on epoch and geography and differed many times from the *theodiscus*, the people's language. Thus, only people capable to communicate in the scientific language were able to participate in the discourse. Aristoteles created the nomenclature of practical science containing, f.i. politics and ethics, theoretical science, mathematics and theology and poietic science, including medicine and poetry. Elaborations meant for wider consumption were called exoteric, whereas those works targeting the circle of like brethren *esotheric*.

Methodology and reproducible results did not play the crucial role as they do today. Alchemy, occultism, and religion where all closely related disciplines and influenced what would become modern science. None of these areas is known for a deep methodological foundation and for good reasons: to believe rather than to know was an integral part of a scientific approach back in these days.

In his seminal work *Saggiatore*, published 1623, Galileo Galilei argued to understand nature requires understanding mathematics, otherwise the inner workings of nature would remain unintelligible. He also dismissed both Alchemy and Astrology as incapable to describe nature, a view Francis Bacon already shared 1597 in his essays: To master nature requires to understand nature. Bacons notion of understanding was freed from influential idols of its time like the Greek philosophers Platon and Aristoteles and, with Bacon's words, their *illusions*. However, even a generation later, science was still deeply embedded into religion, occultism and alchemy. Isaac

Newton described fundamental insights in the domain of optics, dynamics, mathematics, and chemistry, using a systematic, methodological approach. When we make use of the adage *standing on the shoulders of giants*, Newton is certainly at the very base of that pyramid. Lesser known is Newton's role as an alchemist. Three hundred sixty-nine of his personal books deal with mathematics and physics, whereas a stunning 170 books make reference to the Kabbala or Rosicrucianism to support his endeavour to find the philosopher's stone. So even Newton still believed in the unity of science, religion, and occultism.

In 1661, Robert Boyle published the book *The Sceptical Chymist,* 1 year after he and 11 further fellows founded the Royal Society. He called for experimental rigor and for describing chemical experiments in a way that others would be able to repeat and verify results. Robert Boyle and the many to follow him in spirit established the mental model of science as a white collar working activity, producing results with a small community, unintelligible to the people. Modern science, a science solidified in methodology, empirical evidence, and reproducibility of results dates back to the founding fathers of the Royal Society.

Over the years, the methodological aspect of conducting research increasingly gained traction, leaving the aspect of reproducibility behind. This changed due to an infamous Excel mistake, which happened to Harvard University economists Carmen Reinhart and Kenneth Rogoff in 2010, to erroneously conclude a significant correlation between high government debt and slow economic growth (Reinhart and Rogoff 2010). The model they employed in their research paper was grounded in theory, yet their results where irreproducible by others, due to not releasing their research data. As an increasing number of economists expressed disbelief in their findings, they finally published the Excel file they based their investigations on. Soon afterwards, other researchers identified that five rows were left out from a formula, which was used to support their argument. However, the damage was done and it is partly to this paper that Europe now experiences an era of government austerity as many statesmen took reference to it. This poses the question of what is more important to the scientific discourse: Methodological soundness or reproducibility of results through availability of data? While reproducibility is a defining feature of research, the extent to which it should characterize it is debated (Nosek 2015). It can be noted that newer movements, in particular in relation to scientific computing or computational social science, with the increasing importance of big data research, social network data, and machine-generated hypotheses (Lazer et al. 2009), emphasise the importance of reproducibility; in particular since there have been claims of its absence in some domains (f.i. in the area of psychology, where research subjects are rarely static). "In short, a computational social science is emerging that leverages the capacity to collect and analyse data with an unprecedented breadth and depth and scale." (Lazer et al. 2009, p. 722). Computation often reaches into traditionally qualitative fields, also in the area of dissemination, where data sharing and open standards are emerging, and sometimes endorsing pre-publication and open science on the complete research spectrum. Another popular example is Diederik Stapel, a professor of social psychology, who could not produce the data behind his work until he admitted in 2011 that he had been fabricating the data. Apart from these more extreme examples, a scientific movement called reproducibility movement has been formed, and the community pushes not only for publication and sharing of data, but also for the possibility to reproduce results. While irreproducible evidence does not mean that results are wrong, it could also refer to undetected variables.

In his highly disputed book *Against Method*, Paul Feyerabend claims that the idea of a method that contains firm, unchanging, and absolutely binding principles for conducting science meets considerable difficulty, when confronted with the results of historical research. There is not a single rule, however plausible, and however firmly grounded in epistemology that is not violated at some time or another. He claims that such violations are necessary for progress (Feyerabend and Hacking 2010). This and many more propositions discussed by Feyerabend bear lots of controversy, as they are shaking on the still young pillars of what just became "traditional" science.

Neglecting the discussion onto which more attention should be laid upon – the availability of scientific data or a sound methodological approach – there seems to be agreement that scientific research should become tangible for many more people than it is today. Furthermore, we observe a shift towards research impact, visible in the increasing importance of quantitative research measures and automatized citation indexes, like Google Scholar for impact monitoring[1] (Harzing and van der Wal 2008). With an increasing amount of people becoming part of the *scientific community*, a term, which constitutes no sharply-delineated area anyhow, new ways of how to conduct research are emerging.

Open Science

How science emerged and was conducted changed significantly over the past centuries and is still undergoing rapid shifts and changes today. In former times, scientific activities were rather performed by the aristocratic society than by common people, as the trustworthiness of the associated results was strongly interconnected with the scientist being a "gentleman". Yet the situation has changed more and more in favour of repeatability and availability of data than relying purely on big names and the reputation of huge organizations. While science has sought to include outside expertise (Carpenter 2001), the view on the notion of the expert itself also underwent a significant shift. Taleb notes that a great deal of important scientific discoveries with significant impact did not result from planning and foresight, but mostly resulted from a trial and error approach and the unexpected (Taleb 2007).

With view to the inclusion of expertise in ideation systems, different approaches to include outside knowledge or expertise have been classified, mostly focusing on

[1] Also in the e-government or e-policy domain, cp. f.i. Scholl, H.-J. (2016), Profiling the Academic Domain of Digital Democracy and Government, presentation at CeDEM16, conference for e-democracy and open government, 18th May 2016, Krems, Austria.

a top-down approach. Management theory distinguishes between flat or hierarchical forms of including outside perspectives: While the closed “elite cycle” is a more traditional way of production mostly lead by public institutions, other models like the “consortium” are based on a flat governance structure, but still focusing on closed participation. Between the closed hierarchical model and an open-model, communities of practice or creation have been proposed (Sawhney and Prandelli 2000). In particular with view to increased open research data output, community innovation could be fostered in the research context, focusing on the role of communities or crowds, networks, and less hierarchical structures (Parycek et al. 2016). Methods such as crowdsourcing and crowd-based initiatives can be seen as a way to use collective intelligence for innovation. Research further separates crowds and communities, which are distinguished by a set of organizing principles and by “light or heavy-weight models of peer-production” (Haythornthwaite 2009). An example would be Wikipedia, which is mainly crowdsourced, yet also contains structural aspects of communities. With view to citizen science, different levels of engagement, involvement and participation are distinguished, which will be addressed later in this chapter and related to the ePolicy cycle. It can be estimated that with increased experience in network structures and crowds, institutions such as governments and universities will gain more flexibility in utilizing the principles of the network society and opening up their processes on different stages of the cycle.

The open paradigm has certainly found its way into science, next to a counter movement of closed pay journals with other paradigms and goals. Looking at data as one important element and basis of scientific output, the increase of open data output in research as part of the open science concept is recently much supported by the European Union. This is visible in efforts to make the results of publicly funded research freely available within the next few years, as Competitiveness Council agreed on the target year 2020.[2] These changes are part of a set of recommendations including improved access to and storage of research data. The next step in such endeavours would be to enhance the value of open data by increasing activities to transfer it into knowledge and to foster further evidence-building by its usage.

Friesike et al. (2015) extract the main streams within open science and define the following four perspectives:

1. *Philanthropic perspective*: Until recently, scientific knowledge and outputs, paired with the required tools and infrastructure were restricted to a particular group. Yet, universities and research institutions are opening their courses and curricula to public audiences via f.i. downloads or video streaming services such as YouTube. In addition, the advance of open access journals distribute scientific contents to everybody interested in the research.
2. *Reflationary perspective*: Another trend is the publication of intermediate work results in form of pre-prints or even before submission. This approach supports

[2] Enserink, M., In dramatic statement, European leaders call for “immediate” open access to all scientific papers by 2020. Science, 27th May 2016, http://www.sciencemag.org/news/2016/05/dramatic-statement-european-leaders-call-immediate-open-access-all-scientific-papers (accessed 15th July 2016).

researchers in reflecting on their initial thoughts, while at the same time promoting new ideas within the scientific community and beyond; even influence entire research directions in the long run. These published ideas can be commented, evaluated, or even challenged by other scientists or amateurs. Furthermore, the initial starting point of a concept and its evolution over time can be traced more easily this way, as the pre-published versions stay within the Internet even after the final paper has been accepted and published by a publisher.

3. *Constructivistic perspective*: Arising co-creational processes open up new ways of publication development. This includes new and innovative business models as well as associated user models. A prominent example for such an approach is crowdsourcing in which the wisdom of the crowd is used solve problems in a fast and flexible manner and citizens are required to support professional scientists' work, but raising scientific issues or drawing upon problem-solving strategies are still done by professional scientists (Dickel and Franzen 2016). Open platforms with small groups of experts loosely moderated and support the discussion and dialogue between involved parties. But not only problem-solving but also data collection are part of these perspective.
4. *Exploitative perspective*: This perspective refers to real life applications and application-orientated knowledge exploitation in cooperation with practitioners.

Citizen Science

Finke notes that the English term "citizen science" is related to a predominance of the Anglo-Saxon countries in this research area. However, with view to the actual content, he constitutes no big national or cultural differences (Finke 2014, p. 37): everywhere people participate on the collective acquisition of knowledge and on forms of knowledge transfer. While his claim that scientific engagement is not based on profession, titles or control structures, but on interest, skills and activities can be debated, it seems obvious that citizen science can only be realized on the basis of such attitudes. For Finke, the term of the amateur or layman is significant for citizen science. Rationality (German: "Laienrationalität") enables citizen science in a continuously more complex world. Citizenship means to be engaged for something. Citizen science according to Finke satirizes a too narrow understanding of a science that is done only by professionals (Finke 2014, p. 40). Irwin defines the term: "Citizen Science" evokes a science which assists the needs and concerns of citizens. He further notes that the term also makes the point for a science that is *developed and enacted by citizens themselves*. (Irwin 1995, p. xi). Feyerabend (1978) even claims that the amateurs are the only citizens that can be trusted to criticize or monitor science independently. Crucial in this regard is that the distinction between citizens and scientists is blurred, and emphasis is put on the context of scientific work: on everyday life and the lifeworlds of citizens. This claim corresponds well with newer theories of citizenship and participation fostered by the affordances of everyday life, hybrid media environments, e.g. the concept of mundane citizenship by

Bakardjieva (2009) or, with reference to functions of monitoring and criticism, to the monitorial citizen as described by Schudson (2000). Consequently, Finke (2014) defines “being close to real life” as a principle of citizen science: everyday life knowledge is situated in the scientific community. Citizen science is science in the lifeworld of the people, whereas professional science decidedly seeks to abstract from it (Finke 2014, p. 65). Citizen science as a situated and bottom-up practice taking into account broad networks of people is also referred to as “extreme citizen science”, taking the participatory element of citizen science to the extreme (Haklay 2010).[3] In this view, participatory science is the consequent next step of citizen science (Stevens et al. 2014).

Newman et al. (2012) provide a comprehensive overview of the overall evolution and current trends regarding the paradigm citizen science, which is summarized in the following.

In the past, people acted mostly on an individual basis and were driven through hobby-level scientific interests. In return, collaborations occurred on a local scale only. The research questions to be pursued were based heavily on a top-down approach. The process of collecting data was performed with the help of protocols designed by experts in paper-based forms and therefore access to these data was very limited in time and space. The analysis of the gathered data was solely performed by scientists, who published their results in scientific publications. The impact caused by the projects was not a focus and therefore was not a major concern at that time. The motivation behind the conducted experiments was most of the time based on individual interests, rooted in personal observations of the environment and was very limited in terms of technological possibility regarding data collection and analysis.

Today, people cooperate on a national and international level via common projects. While the main source for research questions still is top-down, more and more bottom-up methodologies are arising. Some approaches relate these methodologies and the proliferation of citizen science explicitly to the availability of new technologies, e.g. by mobile data submission (mobile applications or online submission forms) or social networking sites.

Data that have been collected in the course of the projects are now kept online, with a particular focus on aspects such as data quality and data integration. In former times, analyses have been available for local micro scales only. Today, analyses for macro scales are available as well. Further-more, additional efforts are put into the investigation of spatio-temporal phenomena. Yet, the core analyses are still performed by scientists. While the results are still published by scientists in most cases, research related data is made available only to be accessed by all involved/interested stakeholders. The evaluation of results is done via key performance indicators and specific to the current project context, which in turn makes it difficult if not impossible to transfer these assessments and often also to compare the results between projects. While the composition of research teams has improved in terms of diversity, demographic data still indicates the need for further developments in this

[3] The Extreme Citizen Science Group at UCL London is also working with marginalized communities in citizen science activities with the goal to enable wider participation by lay people.

regard. The main motivational driver for participation in these projects is based on individual interests regarding collaboration-related social aspects. The technological adoption rate has increased significantly, as online-based citizen science resources such as blogs offer data publicly to be integrated in own projects.

Enablers and Methods of Citizen Engagement

Citizen science has been described as participatory science (Conrad and Hilchey 2010; Carr 2004). While the use of volunteers has always been an important component, it has evolved into citizen science within the past two decades (Catlin-Groves 2012). This can partly be explained by the use of ICTs fostering forms of participatory science.

While some forms of citizen science refer to a more active form of engagement, a good deal of participation in digital late modernity is based on more mundane, implicit, opportunistic or more passive forms of engagement. As Bennett and Segerberg note on the characteristics of contemporary networked societies, a different form of organisational structure enabled by phenomena of connective individualism (Bennett and Segerberg 2012) and expressive issue-engagement (Svensson 2014) emerged. This has mostly been explained by a specific form of collective action, initiated by the personalisation of actions. With this different "logic of connective action" (Bennett and Segerberg 2012) and the ubiquitous utilization of new media and technologies, the structures of mobilisation and techniques for citizen engagement have transformed. The argument has been put forward that communication technologies replace the need for traditional communities of action, in other words: those technologies take over what has traditionally been done by humans, making it easier for humans to organise themselves and to reduce the cost of organization and sharing.

Research has emphasised the importance of civic engagement as the actual strength of citizen science (Finke 2014). This can also be done in a more continuous form. Also monitoring function thus does not have to come at the end of a process, but can be executed permanently. In this context the potential of online media can create a *multitude of responses and reactions* (Papacharissi 2009, p. 230).

While modern citizenship assumes an active role (cp. The term "DIY Citizenship" by Ratto and Boler 2014), not all modes of participation in the digital networked society have to be completely active. Research has also emphasised the importance of less active form of participation, e.g. in the form of so called "lurkers" (Nonnecke and Preece 2000), who are less active and remain silent, but nonetheless are an important factor of online engagement. In an extreme form, citizens can provide their data as sensors. Information can now be packed digitally and travel anywhere in the world. On the basis of this speed of flow of information coupled with its "relative uncensorability" (McNair 2009, p. 223) and the collapse of time-space "distantiation" (Giddens 1990) and the assumption that members of society have access to and can afford to buy the hardware, the sharing of information has become a commonplace of cultural life, leading to a different form of communication with a lot of

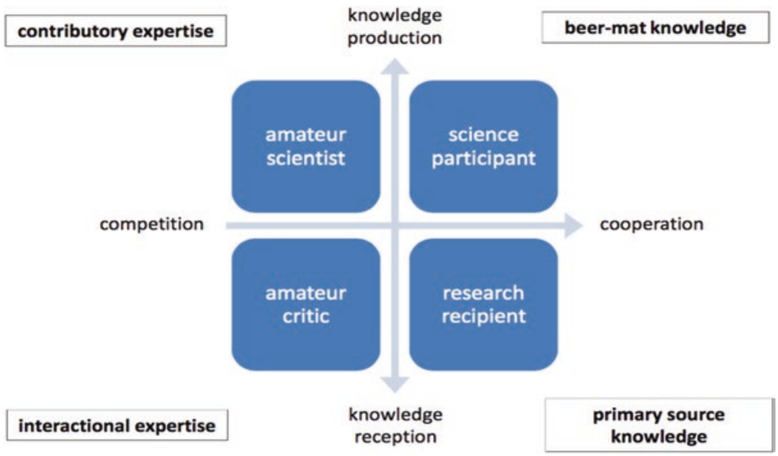

Fig. 1 A framework for engaging expertise, Dickel and Franzen (2016)

data remaining unused. This expanded information flow makes participants constant producers of data, amounting to a globalized public sphere (McNair 2009).

Catlin-Groves distinguishes on the citizen landscape from volunteers, citizen sensors and beyond (Catlin-Groves 2012). In this classification, virtual citizen science refers to data mining in a passive framework (f.i. via social networking sites), which can also have a more active form in the form of active participation. Furthermore, citizen science can comprise "citizen sensing" as an active framework via mobile submissions.[4] Catlin-Groves notes a move "from standardised data collection methods to data mining available datasets", well as the "blurring of the line between citizen science and citizen sensors and the need to further explore online social networks for data collection" (Catlin-Groves 2012). In the context of citizens providing data, (Cooper et al. 2007) emphasise a distinction between "citizen science" and "participatory action research". Citizen science should ideally not use citizens on unequal terms and treat them as scientists on equal terms and not foster a state of competition (cp. Finke 2014).[5] A framework for engaging expertise or knowledge has also been proposed by Dickel and Franzen (2016), who categorize two dimensions in four levels of expertise, which are comparable to science and relevant for policy makers (Fig. 1).

These roles are not found in empirically pure form, but seek to conceptualise inclusion efforts in citizen science. Apart from the differentiation along the needed expertise, these roles distinguish whether the link to the expertise is characterised by competition or cooperation. When characterized as competition, inclusion efforts are expected to be rejected (Dickel and Franzen 2016; Finke 2014), and competition

[4] It can be noted that these newer forms of citizen engagement re less standardized, but mostly opportunistic or directed.

[5] Data compilers should be able to utilize centralized data to produce scientific results in exactly the same way as anyone else should be allowed.

between amateur science and professional science is usually implicit. It can become explicit f.i. when publications of amateur scientists are criticised by the academic world or the other way round (Dickel and Franzen 2016).

Participation in the citizen science landscape can be based on more than intrinsic motivation. The willingness to share can be based on civic engagement, the joy of discovery, but also on more playful motives and play instinct (Finke 2014, p. 124). Another enabler is the private knowledge motives of participants or self-selected areas of interest, sometimes in the form of hobbies and the will to preserve and create knowledge. Behavioural approaches to spatial data sharing have also emphasised the importance of the following contextual factors for the willingness to share: attitude (f.i. strategic position or social outcomes), social pressure (f.i. of institutions, moral norms or the market) and perceived control (f.i. technical or interpersonal skills or finding sharing partners) (de Montalvo 2003).

While those motivational factors play a big enabling role it should also be noted that limited access to technology or technophobia can play a role, and factors explaining motivational access to technology can be of a social/cultural or a mental/psychological kind (Van Dijk 2009). Many technologies do not have appeal for the low-income or low-educated though, and if citizen science is to be appealing to such people, computer anxiety or technophobia as major barriers to access has to be taken into account, as these phenomena are not expected to disappear with the ubiquity of networks in the digital age (Van Dijk 2009). However, technologies of communities (Irwin 2001) make it easier for citizens to participate when they feel like.[6]

Another strategy in lowering the participation threshold is the integration of elements of gamification or game-related elements. Thiel (2016) undertook a meta-analysis of the use of such elements in the field of digital participation. She concludes that while gamification does not work similar in all domains, if situated carefully in the relevant context, gamification could increase the level of participation in some areas and under specific circumstances. However, several studies have already proven that the strategy of adding game elements to influence users' behaviour can be successful: "The most common objective behind gamification is to increase the usage of a system. Other scholars have shown that game elements can increase the perception of effort, make tasks or services more enjoyable and control behaviour." (Thiel 2016, p. 7).[7] Others have found that gamification had no effect in the context of a citizen science application (Bowser et al. 2013): it was found that in an intrinsically motivated user group the game elements in a citizen science application were almost incidental. This can be explained as citizens were intrinsically interested in the non-game context and did not need an additional motivator. Thiel concludes that only if game aspects are utilised correctly and contextualized, they can build a highly motivational user experience (Thiel 2016, p. 8). However, the gamification approach can be effective in terms of influencing or tapping into users' motivation

[6] Irwin explores the configuration of the scientific citizen within policy and consultation processes and accesses the significance of such technologies for the practice of scientific citizenship.

[7] Thiel also addresses that ethical considerations need to be considered.

up to a certain level in order to create a first motivating environment (cp. f.i. on the agenda setting level).

With view to digital infrastructures, methods of science-driven crowdsourcing enabled by the digital are described by Dickel and Franzen (2016), in which a task normally performed by members of an organization is outsourced. Forms of such crowd science relevant in our context also comprise delegating online data collection and assessment to the public. That way, crowd science enables the implementation of large data-intensive projects, which could otherwise hardly be implemented (Franzoni and Sauermann 2014). As Dickel and Franzen (2016) note, knowledge production and the reception of knowledge are becoming increasingly socially inclusive. This raises the question of how much more inclusive new institutions should be and how confidence can be guaranteed if the cycle of experts is expanded. They propose a typology of digitally-supported inclusion models, and on that basis conclude that the line between certified experts and laypeople is blurring (Dickel and Franzen 2016, p. 3).

Big Data as a Technological Enabler

The preceding section primarily dealt with intrinsic factors of motivating participation in citizen science, while this section focuses on extrinsic enablers, with a closer look on big data related technology. We further ask what this could mean for supporting and evaluating governance processes and policy.

It sometimes feels like our society is obsesses with numbers. Scientific theory mostly sees this as a good thing – reproducibility requires prove on the basis of facts, figures numbers. Deming, the inventor of modern quality management and heavy influencer of the reconstruction of post-World War 2 Japan towards the world economic powerhouse of the 1960s, 70s and 80s, coined the following phrase: “In God we trust; all others must bring data”. Books on Amazon with titles referring to data divination are selling well. What does this mean for the future role of the citizen scientist and how does it affects our society? More precisely: How will policy making be conducted in the future? Let’s start with some big numbers first.

Our known universe consists of roughly 10^{80} atoms, a number impossible to fathom. Written out it spells as one-hundred thousand quadrillion vigintillion. Yet Peter Norvig, Director of Research at Google, tends to disagree and argues the (small) number of atoms in the universe. In a blog post referring to Googles breakthrough in beating a human being in the board game of Go, Norvig addresses combinational theory. For example, the number of combinations made possible by a 40-character passphrase, consisting of uppercase, lowercase, numbers and special characters, already reaches the numbers of estimated atoms in our universe. Comparatively, the board game of Go with a 19 by 19 field setup entails 10^{170} legal positions. In other words, combinational theory, which is by nature multiplicative, dwarfs every number of our additive physical nature.[8] Translated to the citizen science domain, in 2015, 3.2 billion people had access to the internet and they are all potentially connected (ITU 2015). This theoretically entails an incredible number of possibilities to share and re-combine data and translate it into valuable knowledge for individuals, business making (What will be the next product a customer buys?) and government (Where is the best place to build a new hospital?)

Combinational theory is just one aspect of the transformational power of ICT enabled by network-connected infrastructure. It is reminiscent of Metcalfe's law, which states that the value of a telecommunications network is proportional to the square of the number of connected users of the system. In other words, every citizen creating data theoretically exponentially increases the value of the network.

The Digitization of Information and a New Breed of Intelligence

Around 2000, two remarkable events related to digitisation took place. First, the amount of digital information surpassed the amount of analogue information. Second, the speed of data and information creation significantly accelerated. Today, a multitude of devices is available at comparatively low costs, enabling the maintenance of networked connections and sensing a multitude of data points; be it RFID-chips, the Internet of things, city sensors or connected sports gadgets. General purpose computers with low power requirements like the Arduino[9] or the Raspberry Pi[10] sell for around 50 € and enable their owners to conceive all sorts of integrated gadgets like home automation devices, weather stations, and beer brewers[11]. However the most widespread digitisation device in use is the smartphone. According to Statista, in 2015 there were 1.8 billion smartphones in use worldwide,[12] which are connected to the Internet most of the time (Fig. 2).

How is this related to citizen science for twenty-first century policy making? Another puzzle piece in our line of argumentation is that of intelligence. When thinking about intelligence, what springs to our mind is human intelligence or secret services. We do not know for sure if people's intelligence changed much in the last three hundred years – the time frame in which modern science formed. Certainly the artefacts we create are increasingly impressive, but this may to a large extent be due to collective intelligence and how we are able to pass knowledge through objects rather than through genes. With view to citizen science something is of greater importance: the ability of algorithms to cope with the plethora of data and informa-

[8] http://norvig.com/atoms.html, retrieved 2016-07-12.

[9] https://www.arduino.cc/

[10] https://www.raspberrypi.org/

[11] http://www.networkworld.com/article/2290609/computers/computers-153240-20-cool-things-you-can-do-with-a-raspberry-pi.html, retrieved 2016-07-12.

[12] http://www.statista.com/statistics/330695/number-of-smartphone-users-worldwide/ (data from eMarketer), retrieved 2016-07-12.

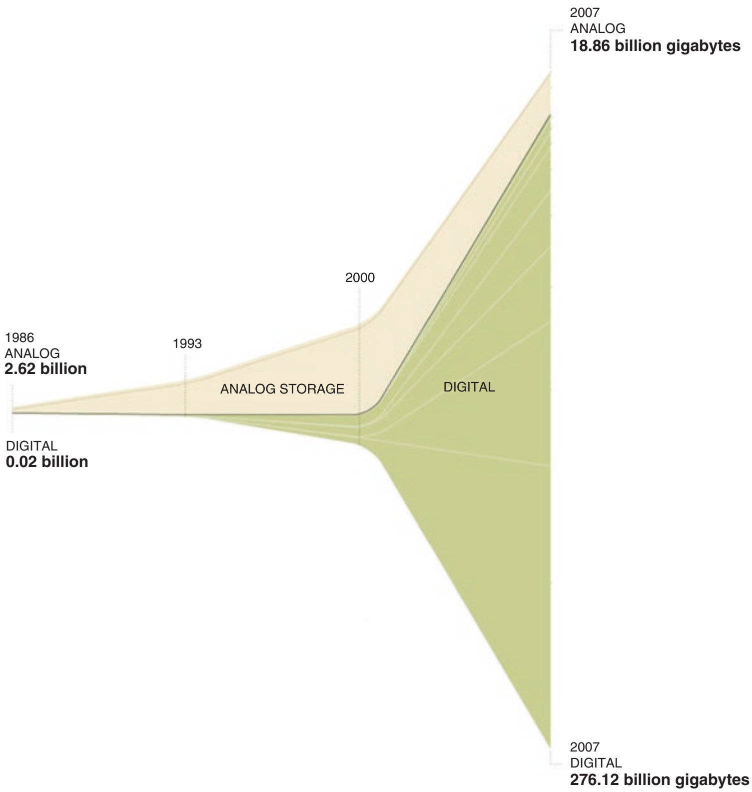

Fig. 2 As of 2000, more information is available in digital rather than analogue and the speed at which data and information accrues tremendously increased (Hilbert and López 2011)

tion generated every day. While the combination of networked devices, exchanging data and information can be the source for better decision-making, it's the algorithms that provide us with the means to actually do so.

Looking back at the combinatorial features we previously identified, the sheer amount of data would be far too large to store, inspect and analyse by any computer system using traditional algorithms. A new way of thinking about problem solving emerged. Striving for optimal solutions in Big Data requires the usage of algorithms which expose polynomial runtime behaviour. Dedicating more computational power in terms of available computing cycles, network speed and storage capacity becomes unfeasible and increasingly impossible. A practical solution outplays optimal solutions which, due to their runtime complexity, may only be able to process a fraction of the available data and thus lead to local optima. "Good-enough" algorithms become necessary if the amount of available data gets too large to be handled by traditional ICT systems (Mayer-Schönberger and Cukier 2013). Imagine an international online retailer. Even such seemingly simple questions such as "How many items of X have we sold today in region Y?" become impossible to answer, given the amount of data accrued over time.

Another crucial aspect of today's ICT systems is the capability to speedily react on external events. This requirement for speed may either be triggered from a single sensor continuously transmitting data, a sensor network whose collectively gathered data results in a continuous data stream, or diverse and heterogeneous data sources combined, like sensor and social media. Instant access to analysis results is paramount.

An illustrative example to this new sort of intelligence we would like to present is the HyperLogLog-Algorithm (Heule et al. 2013). This algorithm on the one hand can deal with enormous amounts of data, yet at the expense of being not 100% accurate. However, this is made up by the ability to analyse many facets in the database to potentially identify multi-perspective patterns. Additionally, this algorithm operates stream oriented, i.e. directly on the data as it arrives at ICT systems. Instead of requiring an additional analysis step, analysis data is available in real time. This is the sort of intelligence we introduced before and which completes the triangle of The Digital Virtuous Forces. It is also this breed to algorithms which prevents misinterpretations in data sets by an ill-chosen or arbitrarily chosen data sampling rate. The importance of correct sampling is well known to statisticians and an integral part of every 101 statistics course. The danger of taking adverse decisions based on incorrect or skewed samples can be adverse to harmful, depending on the consequences drawn from the data. If it's a million dollar business behind, correct sampling becomes paramount. Imagine an online retailer, collecting a vast amount of behavioural data (the "user journey") every day to improve the customer experience and to early react on changes to interaction patterns. Taking no decisions at all can be better at times instead of taking the wrong decision. That's what has happened to Internet giant Ebay in 2003. Back in 2003, Ebay collected a vast amount of web interaction patterns but was only able to analyse parts of that precious data. Future decisions were based on the reliance on correct or good sampling techniques. Analysts knew that due to their inability to incorporate all the data into their decision and alert models, valuable data patterns will remain undiscovered and spurious patterns arouse where there are actually none.[13] Using algorithms, which can inspect the data in its entirety yet at the cost at not arriving at absolutely exact results, was favourable for Ebay.

By describing the changed characteristics of ICT systems we introduced an important concept which we think will change the way government policy is made at each and every level in the future: big data analytics. Big data may be defined as the "cultural, technological and scholarly phenomenon" made up of the interplay of algorithmic analysis of large datasets in order to identify patterns (Boyd and Crawford 2012; Ulbricht 2016).

While the technological dimension is emphasised, it should also be noted that big data also entails an important cultural dimension, in our context referring to the

[13] Cliff Saran: How big data powers the eBay customer journey. Case study, Computer Weekly, 2014-04-29 (http://www.computerweekly.com/news/2240219736/Case-Study-How-big-data-powers-the-eBay-customer-journey, retrieved 2016-12-11).

growing significance and authority of quantified information in public administrations and decision making (Rieder and Simon 2016). Drawing on the thesis that big data is said to advance government efficiency and support evidence-decision making, potential risks and challenges should also be considered. We will briefly cover them in the last chapter.

This section explained the role of ICT to shape the digital citizen sphere and presented some methods to foster citizen engagement. The following section will discuss a big data powered policy cycle including the citizen scientist.

The Big Data Enabled Policy Cycle

The widely accepted model for the design of government policy making is the policy cycle. Originally described 1956 by US political science researcher Harold Dwight Lasswell, the policy cycle provides a theoretical frame to explain government policy making. Depending on the chosen abstraction level and granularity of the step model, (a) Agenda Setting, (b) Policy Discussion, (c) Policy Formulation, (d) Policy Acceptance, (e) Provision of Means, (f) Implementation and (g) Evaluation can be distinguished. The cycle is a helpful instrument for all affected stakeholders like politicians, public administration, NGOs, business entities, and the public when organizing campaigns to respect regulations, or which supportive or enabling ICT instruments can be considered. However, the policy cycle does not come without criticism. First it should be understood as a heuristic which requires tailoring to the actual needs. In practice, the sharply distinguished steps will overlap or certain steps left out altogether (Prozesse—Der Policy-Cycle 2009, p. 110). Everet et al. also identify an overemphasis on the process itself rather than quality or performance (Everett 2003).

Arguably the biggest factor of influence to this approved model is technological change. As we identified, the biggest amount of data today is digital, arrives at high speed and is, due to its plentiful sources, of varying structure. Looking at the traditional policy cycle, the model is iterative, with evaluation happening at the last step. This was justified at times when data was primarily analogue and information a scarce good. However, Big Data methodologies provide the means to inspect massive quantities of data in or near real time, to discover new insights through mining yet undiscovered patterns and to visualize complexities in such ways that actionable results can be immediately derived from Kim et al. (2014). The most problematic aspect of the traditional policy cycle is that evaluation happens as a separate and detached process at the end of the policy making process, which wastes time otherwise available for re-focusing of initiatives or dropping unsuccessful measures altogether. It also does not account for the possibility of a continuous inclusion of evaluation and simulation results to re-assess policies based on evidence (Höchtl et al. 2015) (Fig. 3).

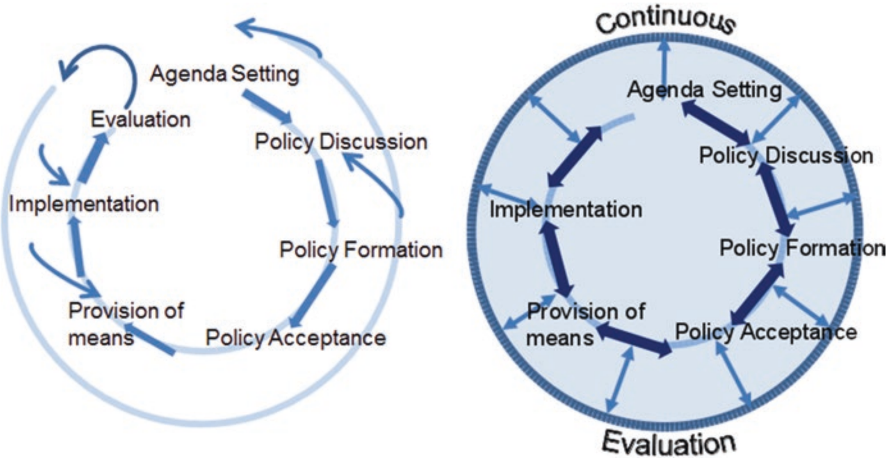

Fig. 3 *Left*: The policy cycle as described by Nachmias and Felbinger, 1982 (Nachmias and Felbinger 1982); *Right*: The big data enabled ePolicy cycle including continuous evaluation

The ePolicy Cycle and the Citizen Scientist

With view to the key concept introduced by Höchtl et al. (2015) of continuous evaluation happening all along the policy cycle, the crucial question is by whom and how evaluation is executed? The administration itself can, will and already does employ big data technologies to better detect tax evasion, forecast disasters based on past damage records, or to address climate change and its effect on the availability of food and water (Mather and Robinson 2016). The tighter integration of yet dispersed data sources is expected to make data based evidence available quicker with the aim to act or foresee large-scale, systemic changes. In the future, algorithms will play an important role in helping policy makers to rectify changes to agreed policies and to instantaneously act on change.

Despite algorithmic approaches, the human ingenuity still excels in detecting patterns in seemingly unrelated data sets. Moreover, citizens increasingly own and operate distributed computing and sensing devices, be it the smartphone or dedicated small scale computers like Arduinos or Raspberries. Therefore the inclusion of citizens into the policy evaluation phase in an organized, structured way including scientific means could draw on citizens' skills, creativity and curiosity for supporting the evaluation of government policy making.

While the inclusion of citizens into government policy making is not new, the ability of citizens to engage in evaluation and monitoring actions in a scientific way is fostered by the availability of big data tools, methodologies and means. However, in the same way as participation will not happen simply by providing the tools and means, incentives and supportive measures will be required to promote citizen participation in science. Depending on intrinsic motivations, personal skills, and interests, a different set of techniques can be employed to encourage citizens to engage in policy evaluation, which may vary from levels of passive participation (lurking), active participation, participation without taking explicit notice (implicit participation) up to coordinated citizen science leagues. Participation enhancing methods such as gamification approaches could also create a breed of citizen scientists without them actually taking notice. The ethical implications of this possibility have to be considered.

Use Cases for the Citizen Scientist in the Policy Cycle

In this section we deduct three use cases of citizen scienceship in policy making, summarise some evidence or enabling elements and analyse the required setup for the successful application of these elements between government and citizens.

Augmented Reality and Gamification Assuming a local authority is undecided whether it should invest in renovating a school or building a new park. There are no legal obligations to prioritize one measure over the other, and even experts are undecided. In a virtual reality environment the government city planners sketch a model of the actual city. People from all around the world subsequently connect to this open playfield and start to model their ideal city. Their activities will become immediately visible to all the other participants of this virtual city. Additionally, every virtual city planner can inspect the planning efforts of the others and what infrastructure he or she has built. After every planning period an election takes place to vote for the chief city planner.

The city has access to the process data of this virtual environment, containing information about which infrastructure was built, which was demolished and how the virtual residents are using their city. They can also see who was elected as chief city planner and replay and analyse the measures taken by her or him. By overlaying the design elements of the virtual city with the actual city by means of augmented reality, the virtual artefacts become immediately tangible.

Enabling elements Assuming that a lot of people enjoy engaging in virtual environments, augmented reality methods for city planning can be successful. One example of a city building simulation in the past is SimCity, which was a huge success even when computers where not yet connected to the Internet. Today peoples' interest in creating an alternate or ideal world has not waned. Minecraft[14] is one example of a game which can be played in a massive multiplayer online mode to design virtual worlds. In *Civic Crafting in Urban Planning*, Mather et al. discuss the potential of using Minecraft for public consultations and argue that serious games in planning can capture participants' attention for a longer period of time, educate the public about planning concepts and site-specific challenges (Mather and Robinson 2016).

[14] https://minecraft.net/en/, available on PC, handhelds and gaming consoles and found its way into many more applications but designing virtual worlds.

Analysis In this use case scenario, the city planning council takes the role of a facilitator by creating a model of the existing city. Additionally it sets the rules to keep people engaged in participating in the virtual planning process, for example by promoting participants to become planning directors, etc. through other players vote. The citizens need not necessarily know that they are taking part in a serious game and that their actions might have an influence in the real world. By choosing a gamification approach, the citizen scientist uses his devices and means to participate, yet the incentives of participation can be "hidden". Instead of scheduling assignments, it is the quest and challenge of the virtual environment which will attract the participants. By using virtual reality elements, the rules of the game can be kept within reasonable constraints, reducing the risk that the citizen scientists create infrastructure which in reality would be inconceivable. The application of augmented reality and gamification to support policy making could be used in the *Agenda Setting* step, where citizens' wishes in the virtual world can be used to prioritize actions in reality.

Ubiquitous computing devices Most smartphone apps fulfil a very specific user need and most users accept trading usability in exchange for granting access to her or his phones sensors (e.g. location) and even more so to contact details. The combination of increased tools usability in conjunction with communicating the goals of the authority could provide another use case. State services would need to provide increased usability levels compared to the offline version or the browser version, e.g. by being seamless integrated into more backend systems without requiring the service user to log into multiple sites to collect information just to enter this information onto another site. Users might then accept the fact that these apps access the phones sensors to deliver data to the authorities, which could support a number of goals, e.g. to reduce traffic jams, or to support early warning systems (rise of temperature in certain regions) in exchange for increased usability. Depending on whether the goal is communicated, users could become citizen sensors knowingly or unknowingly.

Enabling elements The University of Vienna engaged in a joint venture with Samsung to utilize the capacity of smartphones during charging. Cancer and Alzheimer research is computationally intense and involves scanning protein sequences for patterns. Only after the phone is fully charged, a roughly one megabyte large data package will be downloaded by the app Power Sleep,[15] which comes as an alarm clock. The App then inspects and analyses the data package and sends results back to the medical research units.

Analysis The capability to effectively distribute work to many participating nodes in such a way that only little effort is wasted in the coordination of work, combined with algorithms which can efficiently operate on a mere subset of the data, is an achievement of big data research. The citizens' role in the above scenario is that of an active facilitator – he or she will most likely deliberately participate out of altru-

[15] http://www.iflscience.com/technology/new-app-crunches-scientific-data-while-you-sleep/

istic motives. In this role the citizen scientist is unable to influence the details, like the used algorithms, of the performed analysis, which remains under the control of the institution or organisation who is issuing the data for inspection. This is also true for the research results: While the citizen scientist contributes resources, the benefits are harvested elsewhere. The usage of citizen resources by the government is best employed in the *Provision of Means* policy cycle step.

Co-creation Complementary to the voluntary offering of resources by citizen scientists via smartphones in exchange for usability is the idea of planning, designing, and implementing citizens' devices or even infrastructures to sense social and/or environmental phenomena, to collect and aggregate the associated data, and to stream them to a central repository or to provide access to the device/installation via an open API. Such an actively developed networking infrastructure goes beyond the concept of pure data collection and enable participants to actively develop and enhance the underlying scientific ICT infrastructure, transforming the associated projects into living environments. Additionally, the gathered data as well as the research results remain und the control of the.

Enabling elements A prominent example for such an user-implemented sensor network infrastructure can be found in form of the Citizen Weather Observer Program (CWOP),[16] in which private individuals host weather stations that are either using amateur radio or Internet connectivity to transmit collected data. The available sensors range from humidity and temperature sensors, up to sensors for wind speed, barometric pressure and rainfall. While a lot of vendor-sold setups for weather observation exist, a huge group of individuals works with small computerized boards such as the Arduino platform or Raspberry Pies, which provide a high level of extensibility and interconnectivity with other devices and electronic components. Furthermore, the open platforms enable users to freely program their setups in various computer languages. This opens up a plethora of possibilities with view to analytical processes or visualizations.

Analysis Extending the idea to use citizens computing resources, co-creation by citizens requires more intense and ongoing participation levels. Here, a crowd or community of citizen scientists needs to organize themselves, define the objectives, agree on the tools and infrastructure, schedule tasks and governance structures to accomplish a goal. In the most likely case, the government will profit from the results, but seek to secure methodological rigor and soundness of science projects' outcome. The government can support such efforts by legally endowing the opening up of government data and APIs, through specialized research grants also targeting individuals, by providing cloud computing infrastructure which can be used by the citizens like EU's FIWARE platform,[17] or by providing crucial software components as open source like NASA's open source building blocks.[18] Big data tools like platform as a

---

[16] http://wxqa.com/, retrieved 18.07.2016.

[17] https://www.fiware.org/

[18] https://code.nasa.gov/#/

service (PaaS) cloud-computing and cloud-backed decentralized code management services represent technological enablers for citizen science co-creation. Co-creation is best employed in the *Implementation* step of the ePolicy cycle.

Challenges, Issues and Future Implications

Citizen science in combination with big data and evidence-informed decision making raises some issues that should to be addressed at the beginning of projects and throughout the course of scientific investigation (Resnik et al. 2015). In this context, ethical, legal, social and project-related challenges can arise,[19] not only as technology is always situated in a political context (Feenberg 2010), and critical data studies, while in its infancy, have addressed such issues. It seems that all around the world, policy-makers have taken on a hype, and big data is often referred to as the "new oil of the digital age" (European Commission 2012), while at the same time criticised as support of techno-capitalism (Rieder and Simon 2016). Going even further, there is an increasing tendency among citizenry to ignore facts obtained by investigative and data driven journalism. The Trump election campaign or the Brexit were two examples of phenomenon which we might increasingly observe: Neglecting factual proof, irrespective of the efforts and clarity which has been laid on data gathering, model crafting and visualisation making. People believe in what they want to believe.[20] This raises questions of which areas in policy making do make sense to include the citizenry in data driven policy making and to what extend large scale policy making will always remain driven by sentiments rather than by facts, independent of how tangible and easy to understand these facts will ever be presented. This situation is likely to be aggravated by recent advances in non-deterministic and self-improving algorithms like Artificial Intelligence with feedback loops or stacking of algorithms in deep learning arrangements. While the results obtainable by these algorithms or algorithmic arrangements are stunning and are an important aspect to master the complexity of e.g. autonomous vehicles, they are hardly suited for automated decision making, affecting citizens life. Transparency involves many areas such as the availability of data and information for once - the ability to explain citizens why a decision has been made will rise in its importance. The jurisdictions of Germany and Austria have already reacted and grant citizens the right to access the algorithms which have been used to support decision making. This, however, requires the used algorithms to be accessible in a way so their inner working can be explained to the ordinary citizen.[21]

[19] Metcalf and Crawford identified several cases of an "ethics divide" in the big data context and address disputes about human-subjects research ethics in data science.

[20] Down on the Data: facts are not the only truth in life. Greg Jericho, The Guardian, 2016-09-19 (https://www.theguardian.com/commentisfree/2016/sep/19/down-on-the-data-facts-are-not-the-only-truth-in-life, retrieved 2016-12-11).

[21] Data Protection Act Austria (Datenschutzgesetz, DSG), BGBl. I Nr. 165/1999, §49 (3).

While the general consensus is that data analysis can lead to important insights, significant power shifts and advantages and disadvantages for individuals, groups or communities, can arise. Some voices, like cultural critic Slavoj Žižek, have emphasised that humans would not benefit from it, and leaders would probably make decisions not based on data evidence, but still on their own ideological fantasies, claiming that big data analytics would be like “showing Hegel’s logic to a cow”.[22]

Rieder and Simon (2016) argue that while the consequences of big data have been a concern, the *underlying culture of measurement and quantification has not*, and discussions have focused on modalities of change rather than forms of continuity, framed in a narrative of novelty and disruption. Culturally, this can be explained by an effort to reduce uncertainty in societies. The authors address the recent interest in evidence-based policy making and more data-driven forms of governance and relate big data to a distinct political culture based on public distrust and uncertainty. However, more data does not necessarily equal better insights (Rieder and Simon 2016). With the demand for quantitative rigor increasing in societies, a culture of quantification risks reducing the human element, and why the reasons for this shift can be explained as a strategy to adapt to new external pressures, it can also be interpreted as a chance to *de-politicize legislation* (Rieder and Simon 2016). A framework for addressing ethical challenges in citizen science has been provided by Resnik et al. (2015). They propose that for promotion of ethical research, scientists should develop guidelines and provide laymen with education and training on the conduct of research.

Conrad and Hilchey (2010) identified three main areas for challenges regarding the concept of citizen science. While these challenges are situated within their work in the field of community-based monitoring, the authors see them as generic issues regarding the concept of citizen science in general. The first area relates to the *aspect of the number of people involved as well as how to trigger their interest to participate*. This also interrelates to whether or not there exists an established and well-curated network for communication and exchange, which furthermore is also impacted by the provided funding, not only for the citizen science project itself but also for related environmental, organizational, and infrastructural aspects.

The second area covers *challenges in terms of data collection and associated processes*. In order to fulfil many analytical tasks, it is imperative that data are available on a continuous time basis. If the collected data is heavily fragmented, analyses over time become very difficult. Furthermore, there have to be processes defined which provide the necessary means of a guaranteed level of accuracy regarding measurements. Mistakes or measurement errors in the early phase of the project can negatively impact all other succeeding steps. Furthermore, data collected by individuals are always prone to a certain personal bias, and in a more general way, modern data analysis software is often not understandable for the average citizen.

The third area is the *actual use of the data collected within the actual policy/scientific context*, i.e., the adaption by policy-makers in their decision-making process or the publication in a suitable journal. Due to the before-mentioned quality

[22] https://www.youtube.com/watch?v=PBBzYG8szmc (accessed 15th July 2016).

aspects, results are often disregarded as invalid or processes not compatible with the expected level of scientific rigor.

Future citizen science projects have therefore to adapt their processes and overall strategy to overcome these challenges, therefore Newman et al. (2012) foresees future directions of citizen science strongly be based on concepts such as viral marketing, e.g., using social media, interconnected databases, and the initiation of cyberinfrastructures as flexible and scalable backbones. The development of research questions will be predominantly via bottom-up approaches, bringing together practices of amateur research and open science and open source (Dickel and Franzen 2016), supported by intuitive visualization for displaying and navigation data, available in real-time. High quality data will be available 24/7 via globally distributed, high-availability databases. In addition to accessibility, the newly designed cyberinfrastructures offer high-performance, cloud-based computing for everyone, fostering joint collaborations between quantitative and qualitative science fields such as natural and social sciences. The dissemination process will improve due to peer-assessments via social community platforms across the globe. At the same time, this will lead to overall community-accepted key performance indicators, which can be adapted to projects of various scales. The newly formed (virtual) citizen science communities will bridge existing geographical gaps, to enable better and faster exchange and adoption of gained knowledge. The motivation behind participating in these communities will be based on gamification-driven processes, which reward individuals not only with new technological insights but also with reputation within the community, e.g., expressed via achievement badges or ranks.

If citizen science wants to address these challenges, it will be necessary to ask the question how big data relates to power, and how we want to shape the big data society. It is important to note that unethical use of big data can be controlled, and unequal power balances can be recalibrated (Ulbricht 2016). Ulbricht mentions granting wider access to data and data analysis as one way to challenge the privileged position of data collectors and controllers, and also to provide data subjects with participation rights and comprehensible formation. Open data initiatives and increasing public transparency about datasets will be crucial in this context. However, every project should address questions of possible power shifts that might arise, and which unintended consequences they could cause. On the basis of wider knowledge, it will be possible for policy makers to choose the appropriate protection measurements against such threats (Ulbricht 2016). In this context, more empirical studies about the consequences of such projects in the governance field will be necessary in order to be able to make good use of the new instruments.

References

Bakardjieva M (2009) Subactivism: lifeworld and politics in the age of the internet. Inf Soc 25(2):91–104

Bennett WL, Segerberg A (2012) The logic of connective action. Digital media and the personalization of contentious politics. Commun Inf Soc 15(5):739–768. http://dx.doi.org/10.1080/1369118X.2012.670661

Bowser A, Hansen D, Preece J, He Y, Boston C, Hammock J (2013) Gamifying citizen science: a study of two user groups. Proceedings of the companion publication of the 17th ACM conference on computer supported cooperative work & social computing, ACM, pp 137–140

Boyd D, Crawford K (2012) Critical questions for big data. Inf Commun Soc 15(5):662–667. doi: 10.1080/1369118X.2012.678878

Carpenter DP (2001) The forging of bureaucratic autonomy: reputations, networks, and policy innovation in executive agencies, 1862–1928. Princeton University Press, Princeton

Carr AJL (2004) Why do we all need community science. Soc Nat Resour 17:841–849

Catlin-Groves C (2012) The citizen science landscape: from volunteers to citizen sensors and beyond. Int J Zool 2012(2012):349630. http://dx.doi.org/10.1155/2012/349630. http://www.hindawi.com/journals/ijz/2012/349630/. Accessed 15 July 2016

Conrad CC, Hilchey KG (2010) A review of citizen science and community-based environmental monitoring: issues and opportunities. Environ Monit Assess 176(1–4):273–291

Cooper CB, Dickinson J, Phillips T, Bonney R (2007) Citizen science as a tool for conservation in residential ecosystems. Ecol Soc 12(2). http://www.ecologyandsociety.org/vol12/iss2/art11/. Accessed 15 July 2016

de Montalvo UW (2003) In search of rigorous models for policy-oriented research: a behavioural approach to spatial data sharing. URISA J 15(1):19–28

Dickel S, Franzen M (2016) The ‘Problem of Extension’ revisited: new modes of digital participation in science. J Sci Commun 15:1–15

European Commission (2012) From crisis of trust to open governing. Http://europa.eu/rapid/press-release_SPEECH-12-149_en.htm. Accessed 15 July 2016

Everett S (2003) The policy cycle: democratic process or rational paradigm revisited? Aust J Public Adm 62:65–70

Feenberg A (2010) Between reason and experience. Essays in technology and modernity. MIT Press, Cambridge/London

Feyerabend P (1978) Science in a free society. Verso Books, New York

Feyerabend P, Hacking I (2010) Against method. Verso, London/New York

Finke P (2014) Citizen science. Das unterschätzte Wissen der Laien. oekom, München

Franzoni C, Sauermann H (2014) Crowd science: the organisation of scientific research in open collaborative projects. Res Policy 43(1):1–20. doi:10.1016/j.respol.2013.07.005

Friesike S, Widenmayer B, Gassmann O, Schildhauer T (2015) Opening science: towards an agenda of open science in academia and industry. J Technol Transf 40(4):581–601

Giddens A (1990) The consequences of modernity. Polity Press, Cambridge

Haklay M (2010) ‘Extreme’ citizen science, London Citizen Cyberscience Summit, London, 2–3 Sept

Harzing AW, van der Wal R (2008) Google scholar as a new source for citation analysis? Ethics Sci Environ Politics 8(1):61–73

Haythornthwaite C (2009) Crowds and communities: light and heavyweight models of peer production, in proceedings of the 42nd Hawaii international conference of system sciences. IEEE Computer Society, Los Alamitos

Heule S, Nunkesser M, Hall A (2013) HyperLogLog in practice: algorithmic engineering of a state of the art cardinality estimation algorithm. Proceedings of the EDBT 2013 conference, Genoa, 2013

Hilbert M, López P (2011) The world’s technological capacity to store, communicate, and compute information. Science 332:60–65

Höchtl J, Parycek P, Schöllhammer R (2015) Big data in the policy cycle: policy decision making in the digital era. J Organ Comput Electron Commer

Irwin A (1995) Citizen science. In: A study of people, expertise and sustainable development. Routledge, London/New York

Irwin A (2001) Constructing the scientific citizen. Science and democracy in the biosciences. Public Underst Sci 10(1):1–18

ITU: ICT facts and figures – the wold in 2015 (2015) Geneva

Kim G-H, Trimi S, Chung J-H (2014) Big-data applications in the government sector. Commun ACM 57:78–85

Lazer D, Pentland A, Adamic L, Aral S, Barabási A-L, Brewer D, Christakis N, Contractor N, Fowler J, Gutmann M, Jebara T, King G, Macy M, Roy D, Van Alstyne M (2009) Computational social science. Science 323(5915):721–723. doi:10.1126/science.1167742

Mather LW, Robinson P (2016) Civic crafting in urban planning public consultation: exploring Minecraft's potential. Int J E-Plan Res 5:42–58

Mayer-Schönberger V, Cukier K (2013) Big data: a revolution that will transform how we live, work, and think. Eamon Dolan/Houghton Mifflin Harcourt, Boston

McNair B (2009) The internet and the changing global media environment. In: Chadwick A, Howard PN (eds) The Routledge handbook of internet politics. Routeledge, New York, pp 217–229

Nachmias D, Felbinger C (1982) Utilization in the policy cycle: directions for research. Rev Policy Res 2:300–308

Newman G, Wiggins A, Wiggins A, Graham E, Newman S, Crowston K (2012) The future of citizen science: emerging technologies and shifting paradigms. Front Ecol Environ 10(6):298–230

Nonnecke B, Preece J (2000) Lurker demographics: counting the silent. Proceedings of CHI 2000. CHI 2000, ACM, The Hague

Nosek B (2015) Estimating the reproducibility of psychological science. Science 248(6251). doi:10.1126/science.aac4716

Papacharissi Z (2009) The public sphere and beyond. In: Chadwick A, Howard PN (eds) The Routledge handbook of internet politics. Routeledge, New York, pp 230–245

Parycek P, Schöllhammer R, Schoßböck J (2016) Governmental ideation systems. In: Carlsen A, Cerne M, Dysvik A, Skerlavaj M (eds) Capitalizing on creativity at work: fostering the implementation of creative ideas in organizations. Edward Elgar Publishing Inc., Cheltenham, pp 305–319

Prozesse—Der Policy-Cycle (2009) In: Politikfeldanalyse. VS Verlag für Sozialwissenschaften, pp 101–140

Ratto M, Boler M (2014) DIY citizenship. Critical making and social media. MIT Press, Cambridge, MA

Reinhart CM, Rogoff KS (2010) Growth in a time of debt. National Bureau of Economic Research

Resnik DB, Elliot KD, Miller AK (2015) A framework for addressing ethical issues in citizen science. Environ Sci Pol 54:475–481

Rieder G, Simon J (2016) Datatrust: or, the political quest for numerical evidence and the epistemologies of big data. Big Data Soc 2016:1–6. doi:10.1177/2053951716649398

Sawhney M, Prandelli E (2000) Communities of creation: managing distributed innovation in turbulent market. Calif Manag Rev 42(4):24–54

Schudson M (2000) Good citizens and bad history: today's political ideals in historical perspective. Commun Rev 1(4):1–20

Stevens M, Vitos M, Altenbuchler J, Conquest G, Lewis J, Haklay M (2014) Taking participatory citizen science to the extremes, pervasive computing. IEEE CS 13(2):20–29

Svensson J (2014) Political participation on social media platforms in Sweden today: connective individualism, expressive issue-engagement and disciplined updating. Int J Media Cult Politics 10(3):347–354

Taleb NN (2007) The black swan. The impact of the highly improbable. Random House, New York

Thiel S-K (2016) A review of introducing game elements to e-particpation. In: Edelmann N, Parycek P (eds) CeDEM 16. Proceedings of the 6th international conference for e-democracy and open government, IEEE computer society no. P5845, 3-9

Ulbricht L (2016) Big data: big power shifts? Internet Policy Rev 6(1):1–8

Van Dijk JAGM (2009) One Europe, digitally divided. In: Chadwick A, Howard PN (eds) The Routledge handbook of internet politics. Routeledge, New York, pp 288–304

Johann Höchtl is senior researcher at the Centre for E-Governance at Danube University Krems and holds a doctoral degree in Business Informatics. The projects he is involved in include EU-funded research projects and national grants in the domain of large-scale data governance, social media application in administration, open data and inclusive approaches of ICT application in public administration. He is former member of OASIS SET TC standardisation group, member

of OKFN Austria and member of Cooperation Open Government Data Austria, where he is heading data quality sub working group. He was advisor to the E-Georgia strategy for the public administration and advisor to Macedonia (FYROM) to raise interoperability capabilities among federal ministries.

Judith Schossböck is a research fellow at the Centre for E-Governance at Danube University Krems, Austria. Among her research interests are online participation and activism, open government and open access, digital literacy, the sociopolitical effects of ICTs and occasionally cyber utopia or dystopia. She was involved in a study on the internet skills of the 14 years old youth in Austria, the development of a youth participation platform on which four European countries cooperated and a project researching electronic identification in online participation. In 2016, she won a scholarship from the Research Grants Council of Hong Kong (HK PhD Fellowship Award)

Thomas J. Lampoltshammer holds a doctoral degree in Applied Geoinformatics in addition to his Master's degrees in the fields of Information and Communication Technology, Embedded and Intelligent Systems, as well as in Adult Education. He currently works as a Senior Scientist (Post Doc) at the Department for E-Governance and Administration at the Danube University Krems/Austria. Prior to his current position, he has worked as a Researcher and Lecturer in Applied Informatics at the School of Information Technology and Systems Management at the Salzburg University of Applied Sciences Salzburg/Austria. His research experience covers national and EU-funded projects in ICT-related topics, such as Geoinformatics, Semantics, Social Media, Legal Informatics, and E-Health. He is member of the ICA Commission on Cognitive Issues in Geographic Information Visualization and acts as reviewer for several SCI-indexed journals.

Peter Parycek is full professor for eGovernance and head of the department for eGovernance and Administration at the Danube University Krems. He holds a doctoral degree in legal sciences (University Salzburg) and a master's degree in telematics management (Danube University Krems). As head of the department he is responsible for coordinating the eGovernance research groupe, development and coordination of academic programmes and for the management of national and international cooperations with public and private partners. He conducts research at the intersection of legal science, technology and political science. From 2006 until 2011, he was chairman of the Austrian working group e-democracy and e-participation of the Federal Chancellery. From 2010 until 2011, he contributed to the conceptualization of an E-Government-Act for the Principality of Liechtenstein. He has founded the International Conference for e-Democracy and Open Government (CeDEM) and is co-initiator of the open access journal JeDEM