Citizen Science and Public Policy-Making During the COVID-19 Pandemic: Citizen Data Science as an Extended Peer Community

Sine Zambach

Department of Digitalization, Copenhagen Business School, Denmark

Maja Horst

Faculty of Arts, Aarhus University, Denmark

Alan Irwin

The Danish Centre for Studies in Research and Research Policy, Aarhus University and Department of Organization, Copenhagen Business School, Denmark/ai.ioa@cbs.dk

Abstract

Drawing upon the COVID-19 crisis in Denmark, this paper investigates the relationship between public policy-making and Citizen Science as conducted by citizens analysing and visualising COVID data through social media. We consider whether this form of Citizen Data Science (CDS) can help inform public debate and influence policy-making during a crisis. Specifically, the paper explores the CDS community, its contributions and policy interactions.

Though it is difficult to demonstrate a direct causal relationship between CDS and specific policy decisions, we conclude that CDS can act as a form of 'extended peer community' where new developments are reflected upon, discussed and supplemented by a larger and not professionally involved audience. In this way, CDS represents a potentially-valuable input to policy. However, it also raises new questions and challenges for policy-making, not least in terms of policy makers' capacity to identify useful inputs and act upon a broader range of expertise.

Keywords: Citizen Science, Citizen Data Science, COVID-19, Policy-Making, Visualisation, Extended Peer Community

Introduction

Citizen Science has been widely acknowledged as an important source of evidence about social, environmental and natural phenomena: from birdwatching to astronomy and from large-scale databases to local observation (Bonney, 1996; Hecker et al., 2018; Irwin, 1995, 2015; Kasperowski and Kullenberg, 2019; Kimura and Kinchy, 2016). However, and as a recent publication from the Cambridge Centre for Science and Policy emphasized, it is important to consider *Future Direc-*

tions for Citizen Science and Public Policy (Cohen and Doubleday, 2021). Given the attention now being paid to Citizen Science in scientific and institutional terms (Haklay 2022; Irwin, 2018), is it possible that it could make a positive contribution to public debate and, more particularly, decision-making?

In this paper, we focus on the relationship between Citizen Science, policy-making and technical policy advice linked to the COVID-19 pandemic. While policy documents have often been positive about the contribution from Citizen Science (Hecker et al., 2019), it is less clear how exactly policy-making can benefit from such activities – or whether it has the capacity to do so. As Kinchy has noted in a recent review of the literature: "there is no guarantee that decision-makers will take citizen scientists' knowledge claims seriously" (Kinchy, 2024: 218). To put it succinctly: *can Citizen (Data) Science help inform public debate and assist policy-making during a crisis?*

There is a large social scientific literature addressing the relationship between scientific advice and the policy process (see for example Funtowicz and Ravetz, 1993; Hajer, 1995; Irwin, 2008; Jasanoff, 1990). On the one hand, this literature identifies conventional beliefs and myths concerning this relationship: for example, the idea that science and politics can be kept at arm's length from each other, and the recurrent notion that science should speak truth to power. On the other hand, this literature has often explored how science-policy relations operate in practice. Thus, an extensive literature review (SAPEA, 2019) considered the practical complexity of science-policy interactions, the existence of significant uncertainties and disagreements, and the challenges of maintaining a separation between science and politics within pressing matters of public policy. The COVID crisis vividly illustrates many of these points with experts of different kinds being called upon to give policy advice in a context which could be indeterminate and fast-changing, and where the line between science and politics was at times extremely blurred (Evans, 2022; Hilgartner, 2024; Jasanoff et al., 2021).

In what follows, we will investigate how Citizen Science might function as a form of policy input. During the COVID-19 pandemic in Denmark, we observed that citizens were publishing data analysis and visualisations on social media, and that these contributions were being commented upon and possibly utilised by policy makers. Inspired by the rapid growth in data science as a research field, we chose to term the empirical phenomenon we were observing 'Citizen Data Science' (CDS) (Zambach, 2021) and to conduct an inductive qualitative investigation of how it might impact and potentially benefit policy making.

As we will describe in the following section, we use a definition of Citizen Science which allows for a *bottom-up* and *flexible* interpretation (Van Oudheusden and Abe, 2021). In this, we are in alignment with Van Oudheusden et al.'s call (2024) to explore the emancipatory potential of a plurality of citizen sciences by considering it as a many-facetted phenomenon with diverse types of impact. In this perspective, the pandemic can be seen as a form of 'natural' - or 'unintended' - experiment in Citizen Science-public policy relations which deserves to be reflected upon with appropriate lessons identified. In a manner analogous to previous STS discussions around public engagement (Irwin, 2006), this is also part of a larger move to address Citizen Science in more 'pragmatic' (Van Oudheusden et al., 2024), open and empirical terms without reducing it normatively to only 'the good, the bad and the perfect' (Irwin et al., 2013).

Citizen Data Science

Citizen Science has been defined as "science which assists the needs and concerns of citizens … At the same time, [it] implies a form of science developed and enacted by citizens themselves … generated outside of formal institutions" (Irwin, 1995: xi). More recently, Haklay has proposed a four-way classification of Citizen Science activities: crowdsourcing, distributed intelligence, participatory science, and 'extreme' Citizen Science (Haklay, 2013). In line with our previous discussion, Haklay has also emphasized that Citizen Science represents a pluralistic activity, incorporating 'a spectrum of activities and practices' (Haklay, 2022).

The COVID-19 pandemic led to a large international scientific effort (Birkin, 2021; Griffith et al., 2020; Gundelund and Skov, 2021; UKEOF, 2021). It also provided an impetus for the form of Citizen

Science proposed by Irwin (1995), corresponding approximately to 'participatory science' in Haklay's terms, in which citizens themselves take part in processes of knowledge creation (Boisson, 2023). Starting with the original pandemic spread in Asia, there has been much underground or grass root activity in which citizens have run initiatives supporting or informing authorities and populations alike (Van Oudheusden, 2021; Van Oudheusden and Abe, 2021). Forms of this have included visualisation of the regional development of the disease, the development of predictive models and the discussion of behavioural patterns.

Social media have provided a crucial platform for citizens to share knowledge about COVID-19: including infection rates, vaccination uptake and mutations. As in other areas (Sprague and Tory, 2012; Trajkova et al., 2020), subgroups of users have employed this platform to create customized data visualisations on various topics related to the pandemic (Pueyo, 2020; Roberts, 2020). The present paper represents an attempt to focus on the manner in which data collection, modelling and visualisation were performed on social media by people who are not professionals in the area, who are not paid to conduct such work and who do not represent any organisation. We term such people 'citizen data scientists' (CDSs).

Haklay (2022) has discussed related phenomena in terms of 'citizen cyberscience' (see also Grey, 2009). Our treatment here overlaps with that definition – especially regarding its 'community science' dimension. However, in what follows we have chosen to address questions of data analysis and visualisation rather than the larger phenomenon of the use by citizen scientists of computers and the internet. Such activities fit within a pluralist definition of Citizen Science since they involve scientific methods of data analysis, but participants do not do so in any formal research capacity or as part of their main employment.

What originally caught our attention was that this group of CDSs made use of data analysis and forms of visualisation to support critical discussions about policy-making regarding the pandemic in a form which, in structure and content, was similar to advice from recognized scientific sources (for example, university experts). The visualisations from the Danish COVID-19 CDSs consisted of charts illustrating the development of cases, deaths and hospitalizations, as well as vaccine roll-outs and more detailed information, such as regional patterns. These were shared on social media, particularly on Twitter (now X) (Sleigh et al., 2021), which in Denmark at that time was used mostly by elite groups to debate current events. In this account, we do not seek to explore the relationship between CDS and data science, although it would certainly be useful to conduct a study of how Citizen Science has influenced scientific development in that field. Our focus instead is very much on the CDS-public policy interface.

In general, the CDSs on whom we focus did not collect primary data – and in that way are unlike other forms of Citizen Science which have tried to influence policy-making (Dosemagen et al., 2022; Groom et al., 2019: Ottinger, 2021). Rather, they utilize subsets of open data available from multiple sources such as Our World in Data, Johns Hopkins Coronavirus Resource Center and the Danish health authorities' SSI dashboard (Our World in Data; COVID-19 Map; SSI Dashboard). Before we move to our methods section, we will provide a brief overview of the COVID-19 pandemic in Denmark.

COVID-19 development in Denmark

On March 11th 2020, the Danish Prime Minister solemnly announced a comprehensive lockdown of the entire country with only essential activities, such as grocery stores and health services to be left open (Christensen et al., 2021). The public was advised to avoid face to face contact outside of households. No official curfew was imposed, but a lockdown affected most workplaces in the country. This policy differed from for instance Sweden which had a different, network-based approach, and where there were fewer restrictions, no lockdown, and schools and nursery homes were never closed during the pandemic period (Christensen et al., 2023; Nielsen and Lindvall, 2021).

The initial lockdown in Denmark was generally considered successful in terms of controlling the pandemic and from April COVID incidences began to fall. Consequently, the country gradually became more open in May and June. However,

in the Autumn of 2020, COVID incidences started to rise and during the subsequent Winter lockdown incidences were much higher than in the initial wave. Restrictions were gradually eased again during Spring 2021. Vaccines from Pfizer, Moderna, Astra Zeneca and later Johnson & Johnson were approved in the European Union, and in Denmark vaccinations commenced in January 2021. Generally, Danish media and public debate devoted a lot of attention to the number of available doses and the potential dates for vaccination of different age and occupational groups. In April-May 2021 the health authorities decided to remove the vaccines supplied by Astra Zeneca and Johnson & Johnson from the general vaccination programme due to the risk of side effects combined with the existence of effective alternatives. This decision gave rise to some public discussion as many people questioned whether it was justified (Christensen et al., 2023).

In Autumn 2021, there was renewed focus in all media on new COVID-19 variants and a short lockdown was enforced just before Christmas. However, since hospitals did not become overburdened during the Winter, Denmark was one of the first countries to ease restrictions early in 2022 (Forthun et al., 2024).

Method

As already noted, our focus here is on how CDS has contributed to public debate and policy-making. We therefore chose to conduct semi-structured interviews with CDSs themselves as well as professional science advisors and policy-makers. Importantly, the first author of this paper had been following and engaging with the CDS community through several social media during the pandemic. While this active engagement was not originally planned as participant observation, it helped us identify key informants with whom to engage. Building on this knowledge of the field, we employed a form of snowballing (Kvale, 1994) where we started with those who were continuously and actively participating in the debate on Twitter from the hashtag #covid19dk, and epidemiology experts frequently used by Danish media. In addition, we contacted people who were interacting with and interpreting models and visualisations made by the CDSs.

Table 1. List of interviewees (categorized based on current employment).

Reference	Gender	Employment
CDS1	Male	Retired Chemistry scientist
CDS2	Male	Marketing consultant and Pharmacist
CDS3	Male	Economist
CDS4	Male	Pharmaceutical Chemist
CDS5	Male	Computer Scientist
CDS6	Male	Microbiologist
CDS7	Male	Professor in Economics
CDS8	Male	Associate professor in social science
CDS9	Male	Chief medical doctor
CDS10	Female	IT consultant
CDS11	Male	Civil servant
CDS12	Female	Legal consultant
CDS13	Female	Biologist
CDS14	Male	Computer scientist
DJ1	Male	Journalist
EXP1	Female	Epidemiologist
EXP2	Male	Senior Researcher in Computer Science, member of a modelling group
EXP3	Female	Project Leader, a modelling group
EXP4	Male	PhD fellow in Computer Science, member of a modelling group
PM1	Female	Civil servant
PM2	Male	Civil servant
PM3	Male	Member of Danish parliament
PM4	Male	Head of a Danish health agency
PM5	Male	Former head of Danish health agency

In total, we conducted 24 interviews. Interviewees can roughly be divided into three groups. The first consists of *those who had previously shared visualisations or mathematical models of COVID-19* on Twitter (our core notion of CDS) (n=14). The second group consists of *those who had worked professionally with disease modelling* (n= 4). The third group included *those who were part of the policy making process* (n=5). We also interviewed one person working as a data journalist.

These specific groups were selected in order to build up a broad account of the CDS work and whether it had an impact on policy and public debate. We attempted to contact a range of CDSs in terms of opinions and activities. We specifically tried to locate both male and female CDSs but managed to engage with only three women, of whom only one actively created visualisations herself, though all used the publicly available dashboard tools that were released in mid-2020. A list of the respondents, their job titles and relation to Citizen Data Science can be found in Table 1.

For the purposes of this study, we sought to exclude participants with arguments and viewpoints akin to conspiracy theories: for instance, that COVID-19 does not exist or that Bill Gates had added microchips to the vaccines. In general, those in this category did not directly engage with or share the visualisations, official data and mathematical models which were central to our study. Of course, knowledge was developing fast and opinions shifted across the period in question. One example of such shifting interpretations was the volatile debate over the benefits of vaccines, where official policy originally focused on herd immunity, but public critiques centred on individual benefits and risks to people in certain age groups. The latter perspective later became included in official policy, when two of the available vaccines were withdrawn from the Danish vaccination programme.

In making the inclusion/exclusion demarcation, we specifically aimed to include informants whom we assessed to be making data-related arguments that were likely to be accepted by policy makers as relevant. If in doubt, we erred on the side of excluding informants. This is not because we consider more radical argumentation irrelevant to public debate, but because we focused on the form of CDS which would be easiest for authorities to recognize (see also Ottinger, 2021) – and this form of impact is our focus in this paper. However, we recognise that by making this distinction, we are also reinforcing a separation between voices that might be heard in policy-making and voices that are not – a separation which we have considered in other contexts (Horst and Irwin, 2010; Irwin, 2006)

All interviews were video recorded. They followed a semi-structured interview guide (Table 2) and were later transcribed and coded in two steps following Miles et al.'s (2014) process of coding inspired by grounded theory. Discrepancies in coding were discussed among the authors and resolved through dialogue. Coding focused firstly on descriptive information on members of the CDS community, their motivation and their impact on public debate and the policy process. Secondly, we inductively identified four aggregated themes which were seen to be important for understanding impact on the policy process. These four themes will structure our presentation of results in the next section:

1. Description of CDS and the community
2. Motivation of the citizen data scientists
3. CDS impact on public debate
4. CDS impact on the policy process

Our informants did not always agree with each other. For that reason, we aim to represent the diversity of viewpoints rather than one coherent interpretation. Nevertheless, the main themes of our analysis are dimensions that all informants touch upon and where their assessments are relatively aligned.

Analysis

The next sections will first of all describe the CDS community and, following this, discuss CDSs' motivation to engage with COVID-19 data and participate in public debate. In the third and fourth parts, we investigate how our respondents consider and reflect on the impact of CDS work.

Table 2. Interview guidelines. We have included the following codes in the third column: CDS descriptions, CDS Impact and CDS motivation, written in italics. CDS refers to Citizen Data Science, while PM/Pro denotes Policy-maker/Person who worked with epidemic data as a profession during the COVID-19 pandemic. All interviews were conducted in Danish.

Question	Target	Rationale/*code*
Have you noted a larger effort of citizens visualising COVID-19 data? Where? Twitter, Instagram, …	All	To understand where debates have been taking place/*CDS Description*
How have you been involved in the many models and visualisations and debates from this?	All	To provide context/*CDS Description*
How are the visualisations or models better/worse than the professionally created models/visualisations?	All	To encourage the respondent to discuss quality/*CDS Description*
What impact do laypeople's visualisations have on the debate? Can you give positive or negative examples?	All	To initiate a discussion about public debate / *CDS Impact*
Does it have an impact on decisions? (I am thinking of e.g., "the green and the red wave", prognoses based on open covid-data, comparisons with other countries, vaccine stats, over-reactions. How do you measure/demonstrate this?)	All	To ensure we get a concrete answer/ *CDS Impact*
What kind of impact has it had? Can you think of the impact? Main influences?	All	To ensure we get a concrete answer/ *CDS Impact*
How has the effort been rewarding to you? How do the critiques of your visualisations affect you personally? What drives you?	CDS	To understand motivation/ *CDS Motivation*
Would you characterise yourself as a citizen scientist?	CDS	To investigate the CDS concept/*CDS Description*
How could your work support governments/decision makers?	CDS	To understand respondents' own perception of influence/*CDS Impact*
How did it affect your work positively and negatively?	PM/Pro	To understand influence – both in positive and negative terms/*CDS Impact*
Do we need citizen data scientists? For what? What role?	PM/Pro	To assess the potential roles CDSs can play /*CDS Impact/CDS Description*

The CDS community and its contributions

Our focus was on people who were conducting data analysis without being professionally linked to the COVID-19 response. However, during our interviews it became clear that it was not always easy to distinguish between CDSs on the one hand, and professional data scientists on the other; or between active CDSs and other citizens. Some of the Twitter users who interacted extensively with data analysis did not create visualisations themselves, but helped refine and develop those made by others. Likewise, some of the data scientists working for government were often not trained in human epidemiology but were involved based on their analytic skills from other areas. We have made a rough categorization of our informants in Table 1, based on their main employment during the interview. However, we can already conclude that the definition of a Citizen Data Scientist overlaps with related categorizations (including professional data scientists and government employees).

The early period of the pandemic, from February 2020 to May 2020, was characterized by a high level of uncertainty. Social media played an important role as a platform for CDS activities from the beginning of the pandemic and even before Denmark had its first COVID-19 patient. During this early phase, Danish authorities were struggling with obtaining and structuring data and the data sources they developed were not publicly available. This was seen as a problem by CDSs and they began to campaign for greater data access.

From May-June 2020 onwards, data became more freely available and many CDSs posted their work as an addition to the official charts. Figure 1 gives an example of such work. When we asked

why our informants use Twitter for publication of their work, they pointed to the community they can link up with: "For me, Twitter is a knowledge platform where you get direct access to expertise" (CDS12). In addition, the openness of the Twitter stream meant that the CDSs could publish their data modelling and visualisations as part of the public debate in Denmark.

Overall, there was a strong sense of collaboration among CDSs and their followers on Twitter through the discussion and constructive criticism they shared with each other. We can interpret this as an example of social media's capacity to foster communities of practice among citizen scientists (Liberatore et al., 2018; Sleigh et al., 2021), albeit a self-assembled community (in contrast to the description by Liberatore and colleagues). Performing collaborative work through constructive criticism was seen by CDSs as a clear strength and can be construed as a form of 'extended peer community' (Funtowicz and Ravetz, 1993). Several CDSs described themselves as an informal group who work together to improve their interpretations of data, relating to topics such as hospitalizations, deaths and incidences. They also reported that they had been discussing data processing and visualisations to give the best description and avoid misinterpretations. This peer review function within the community of CDSs is seen as important for their own development, but also for the quality of data visualisations and modelling.

Self-identified motivations for CDS

Almost all our CDS informants referred to a wish to act against the unknown when asked about their motivation to participate. They did not want to rely on authorities' and experts' interpretations of the situation but felt a need to engage actively. The need to act was often coupled with frustration regarding the existing public debate and communication from authorities, such as the lack of data, poor quality of visualisations and apparent limitations in modelling methods. As one

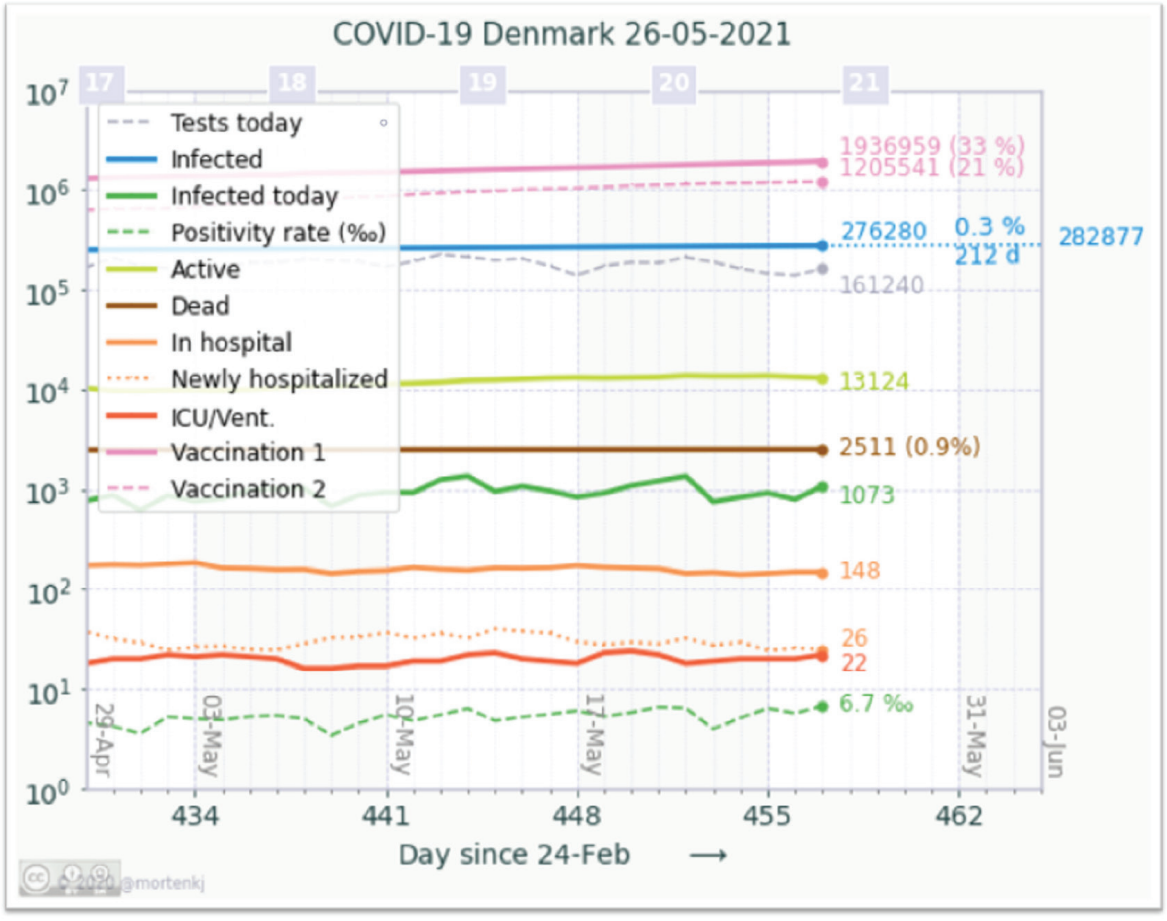

Figure 1. Example of Twitter visualisation

respondent expressed it when reflecting on the early days of the pandemic:

> So I could see that this thing, it does not look good. The data I was looking into, in my world at least, they gave a completely different story than what was generally told in Denmark at the time. And then I started reacting. I started writing to newspapers, and I also started posting some things on Facebook. (CDS4)

Some of our respondents felt an obligation to get involved since they had expertise in epidemiology, economics, data modelling and/or visualisation. We asked all of them whether and how their contributions were different from other voices in the general public debate. Most of them answered in terms of credibility and authority. Since their public views and arguments were based on data, they argued that they had more persuasive effect and represented more important voices: "When you show up with data, it is somehow more supported than if you just have an ordinary public debate" (CDS4).

We also found more personal motivations: "At the beginning I actually had a bit of a competition with TV2 [a Danish TV station] about who was first with the numbers. It's something like that which drives me too - that is, the competitive element …" (CDS5). Working with data also represented a form of emotional check for participants: "I get overview and control over the situation by understanding what is going on" (CDS8). Another was more direct: "it keeps my anxiety down to follow everything that was going on" (CDS12). For one of the CDSs this was the most important motivational factor: "If I am to be completely honest, I will say that it is the psychological part which has motivated me the most" (CDS8). This person continued by explaining why he circulated information on Twitter:

> I think that has just become the natural way to do this. You share your thoughts on social media – these are my thoughts on corona, what do you think? And then it is interesting to see if people agree or disagree and if there is something I have overlooked ... it is a way in which I try to establish control over the situation. (CDS8)

We can gather from this respondent that reactions from other CDSs were an important means of calibrating interpretations of the unknown. The motivation they describe is not simply to act, but to do it in collaboration with others. Generally, our analysis reveals that the CDSs see themselves as motivated by both external and internal considerations. They felt they had knowledge that they were obliged to share, but they also used their CDS work to help bring their own worries and concerns under control.

Impacting public debate on COVID

Most of our interviewees described how CDSs were informing and contributing to the general public debate about the pandemic:

> There are some individuals who are really talented. They are not only good at collecting data, but also at creating a good overall picture and making it easy to understand. Many different types of data are needed to be able to see and understand the whole situation and get an overview. (CDS10)

This description of 'many different types of data' being necessary to get the full picture was common among our CDS interviews. Very often, the CDS Twitter posts received numerous comments from a variety of people and this contributed to shared opinion forming and error finding.

> Especially those who only focus on a specific area manage to contribute something. Because if you delve into the infection figures, or other figures from COVID-19 such as wastewater or vaccine, then there may well be a lot of noise in the data - or volatilities. But if you get used to the noise or these volatilities - for example, that hospitalisations can increase over the weekend because you don't discharge people during this time, but only on Monday - your interpretation can be better. (CDS3)

These interactions fed into the general public debate, which included politicians and authorities who followed the Twitter feed. In this way CDSs described a form of co-created science communication environment, where they participate and learn from their engagement: "We are all learning from these discussions and visualisations. Both others, and us, who participate" (CDS2).

Generally, CDSs themselves and our other interviewees saw their contribution to public science communication about the pandemic as an important form of impact. At some point, the health minister requested his staff to assess whether they could visualise more in line with a specific example taken from CDS work (PM4). And one of the CDSs pointed out how their communication was superior to that from authorities:

> I think we are making better visualisations [than authorities]. We are not forced to include for instance scientific confidence intervals, so we have more freedom with respect to communication as a private citizen. (CDS2)

However, the CDSs also discussed such matters in a more critical manner. As one of them observed: "Epidemiology has been rather free from politicization previously, but this is just no longer the case" (CDS2). Several of them raised the issue of misinformation:

> Unfortunately, I think that there are people who use them to misinform, even with very nice or very complex visualisations. That has clearly been negative - when suddenly it is about politics, rather than facts. I think that has probably been the most difficult part of this. You suddenly realise that it's not just people who share facts and discuss, but also people who have underlying agendas and political drives, which I'd rather be free of. (CDS4)

Overall, we might speculate that CDSs see themselves as having a function as mediators in the grey zone between information and so-called misinformation. The fact that they could discuss scientific modelling at a complex level, but were acting as private people and therefore not bound by obligations to a scientific or other authority, allowed them greater liberty to traverse this field. However, this also points to a major question about quality control of the input from CDS. CDSs view themselves as helping to identify misinformation, but other actors might well characterize their input in exactly those terms.

CDS and the policy process

CDS participation in national debate not only had an impact on public science communication, but also created effects in the policy process more generally. Health authorities and policy makers discussed how questions posed by CDSs had been relevant for their own development of data analysis, just as they had found CDS charts useful as a supplement to official data analysis. A civil servant from the health authorities explained how they had often checked with experts whether models and prognoses from citizens were correct or whether the questions posed by CDS were valid and relevant. This was particularly important in the beginning of the pandemic:

> Things were a bit more chaotic back then and we did not have much data. Therefore, we wrote the health agency if it was correct [what the CDSs wrote] (…) we do not follow everything daily, so sometimes we are made aware of developments through Twitter. (PM2)

Similarly, the head of a Danish health authority reported that he had learnt a lot from the CDSs and that he respects their work as a form of 'double checking'. He also mentioned an example of a CDS with an economics background, who had made a prognosis for the pandemic that supported the official account although the calculations were a little different. It seems that this form of public review was seen as beneficial by the official model makers:

> I think - and this is also an attitude we have in the model group - that we have nothing against that people go out and make their own models or re-calculate ours. The model calculations are really a strength because we can double check the results and question our models when we do not get the same results. (EXP3)

A Member of Parliament with a scientific background also pointed to public review as an important function of CDS: "I think that sometimes errors by the model group are captured and corrected afterwards because citizen scientists have made them aware of the error" (PM3). However, this form of review is not simply about finding errors. An expert with a central role in the official

modelling work described how the outside input helped prevent premature consensus among the official model-makers:

> The qualified input has especially helped to pull you out of your own thought pattern, which you can get a little trapped in if you work so intensely with a specific agenda. When there are so many things that must be done in such a short time, there is a pretty big risk that you won't be able to stop and think twice […] it can be enormously useful that some qualified people from outside come with constructive criticism. (EXP4)

Nevertheless, some of our respondents from the Danish health authorities observed that this form of double-checking could be challenging – and possibly distracting - in the middle of a pandemic: "there are some obvious misunderstandings and then I can become frustrated" (PM5).

A major focus of discussion within our interviews concerned whether CDS had not only general but also direct influence on policy making. Direct influence is very hard to prove. Nevertheless, many interviewees reported that politicians had used CDS posts to pose questions to the government.

One example relates to the decision by the Danish government to remove two vaccines (manufactured by Johnson & Johnson and Astra Zeneca) on the grounds that the risk of serious side effects was too high relative to the risk of dying from COVID. Many Twitter users with knowledge of economics objected to this. They focused on the risk assessment and argued that people could have different considerations of risk – for instance, getting a vaccine sooner would free them to make other life choices with possible health benefits. In addition, a system of free choice would increase the total number of vaccinated people and thereby benefit society. After a while, the health authorities made the vaccines from Johnson & Johnson and Astra Zeneca available once again, conditional upon a prior medical consultation. As one of our informants put it: "Twitter is a form of barometer, and I am sure politicians follow this barometer" (CDS12).

Another example relates to the choice of test strategy. In the first phase of the pandemic, the test capacity in Denmark was very low. PCR testing was the only available test: so-called 'quick tests' (lateral flow tests) were seen to be unreliable. However, one of our informants described how he felt himself to be part of changing this: "I wrote about quick tests at a time when they were unpopular. But some of my politician followers seems to have pressured the policy-makers such that the quick tests were introduced." (CDS6). From late 2020, Denmark developed a very extensive test strategy and the use of professionally administered lateral flow tests became widespread.

It is indeed difficult to demonstrate that CDSs by themselves changed policy-making. However, we do conclude that CDS work has served an important agenda-setting function. CDSs asked questions and posed data interpretations which led other actors, for instance, the media or parliamentary members, to ask new questions, or consider current pandemic measures in a new light. Certainly, the head of a health authority in charge of pandemic surveillance was very positive about the contributions from the CDSs. After our interview, he posted on Twitter thanking the CDSs who had helped during the pandemic.

Our final finding in this sub-section relates to the working conditions of CDSs as well as their future. From the beginning of the pandemic, data was a very scarce resource. CDSs told us about sitting with printed charts and measuring columns with a ruler to obtain the numbers necessary. Later, the official data were published in PDF format, and people had to copy this into Excel sheets or other table formats to make their own visualisations and models.

Based on these cumbersome experiences, citizens (and then politicians) started putting pressure on authorities to publish data in a form that would provide better conditions for the CDSs. From the end of April 2020, the health authorities created a daily dashboard. Similarly, vaccine data was inaccessible in the beginning, and the CDSs complained about this on Twitter. At that time, the head of the health authority in charge of data had just been appointed and the new responsible person explicitly prioritized making data available through the dashboard and basic data files. This was noticed by many of our CDS informants: "I was actually a bit surprised by this – the authori-

ties actually think about us and are willing to use resources on it (…) – it is really positive" (CDS12).

Public access to data and modelling codes was therefore a constant theme. For some of our interviewees (represented in all participant groups), the availability of data is a question of allowing extended review: "It is important that data and code are open for the public. It is not for criticizing, but for qualifying, and I have pushed forward on authorities to do this" (PM3). Other respondents (mainly CDSs and policymakers) stressed that it is a democratic right to be able to access data as well as the underlying modelling code.

This is an important discussion, since Denmark has one of the best public health data repositories world-wide as well as a generally high-quality data system. But when a new issue, such as COVID, emerges there is still a delay: procedures need to be established for how data is to be extracted, collected, and checked for compliance and quality controlled. As some of our respondents commented, the data needs of the official modelling group and the government were prioritized in the beginning and public availability was not at that point seen to be so crucial. It was only later that the contribution from CDS was considered significant. Nevertheless, our respondents agree on the need to make data available, and we can possibly count this another impact of the work of CDSs during the pandemic.

Discussion

This paper began with one question: Can Citizen (Data) Science help inform public debate and impact policy-making during a crisis?

As already noted, definitive evidence of policy influence is hard to locate. Nevertheless, the CDSs in question do point to several ways in which their work can be seen to have had impact. These include influence over the media and politicians, the improvement of data quality and data presentation, direct contact with policy-makers, helping dispel conspiracy theories and deliberate manipulations, and acting as advocates for more accessible information. Open data and data quality have been a core theme within Citizen Science, both in terms of data produced by citizen scientists themselves and, as in this case, making data and coding available and useful for citizens (Bowser et al., 2020; Hecker et al., 2019; Murray et al., 2021; Roman et al., 2021).

More generally, the activities of these CDSs can be seen as what we have termed (drawing upon Funtowicz and Ravetz, 1993 in particular) an extended peer community where new developments are reflected upon, discussed and analysed by a larger audience. This has performed a function of external scrutiny. Beyond that, it has increased the accountability and transparency of decision processes, and – at least at times – helped set the agenda for policy-makers. At a basic level, the Twitter contributions of CDSs served as an indicator of the controversiality or potentially-problematic character of the latest governmental announcements regarding COVID: a barometer of public opinion.

In noting these forms of influence, it is also important to point to a number of ways in which the activities of CDS raise new questions, and new challenges, for public policy-making. One significant aspect of this concerns the relationship between Citizen Data Science, the policy process and science communication. Sitting at the interface between providing technical advice and disseminating knowledge about COVID to a larger population, the CDSs effectively challenge the conventional idea that policy in such a field can operate at a distance from public and political debate. One prime example here was the official decision to suspend the Astra Zeneca and Johnson & Johnson vaccines based on the perceived risk/benefit ratio, only to decide that there were other factors at work: not least other assessments of risk/benefit and the broader societal advantages of more extensive vaccination. In this, the CDSs were both opening up the discussion of COVID policies to a larger audience and at the same time reminding policy-makers that public understanding and public trust are essential if policies are to take practical effect.

Building on that point, we have noted that CDSs have both worked in collaboration with policy makers and, at times, in some tension. The fact that Citizen Science can perform multiple roles in a case such as this - from serving as a barometer of public opinion to agenda-setting - can also make such evidence hard for policy makers

to digest under considerable time pressures. It is not difficult to imagine that civil servants and others could find it frustrating to deal with resourceful citizen scientists during a major crisis but also struggle to set boundaries between the rather focused group considered here and more radical perspectives. This latter discussion could also be found among the CDSs themselves who often sought to distinguish between their own 'informed' views and those who were merely critical.

Whilst this is understandable, the activities of the CDSs could be viewed as a very necessary, and very immediate, form of feedback from a group of attentive and knowledgeable observers. The CDSs served to impress upon policy makers that there are always more questions to be asked and that policies need to make sense in society (and not just in the policy room: Webster, 2007). Though this study presents only the Danish situation, we would argue that this interaction and feedback could be utilised in other countries with some degree of free speech, particularly when a crisis hits, be it a pandemic, the effects of climate change or natural disasters. In that sense, the existence of crisis also represents an opportunity for Citizen Science to help address immediate matters of concern when resources (scientific, institutional and immediately practical) may be in short supply.

An immediate conclusion from our study is that, at least in this case, Citizen Science can be considered a significant asset for challenging and assessing science communication from authorities and for testing and improving public policy making. In that manner, it can offer a form of advice which fits within current understandings of the scientific advisory process but also extends it in constructive ways. However, Citizen Science also offers new questions and challenges.

Among these questions and challenges we would highlight, firstly, the challenge of making policy in a world where external observers can quickly focus on areas of inconsistency, discrepancy and weakness. From our perspective, this extended peer community has many advantages, but it does not necessarily make it easy for the policy maker to operate in an increasingly open system (Haklay, 2021).

The second linked challenge concerns the capacity of the policy process to absorb and respond to the inputs of citizen scientists. In this case, it probably helped that Denmark is a small and still rather consensual country (Horst and Irwin, 2010). Nevertheless, conventional policy processes are geared to expertise of more orthodox kinds and not that which is primarily communicated through social media. Whilst some direct influence can be pointed to in this case, there is also a clear risk that the expertise possessed by citizen scientists can get lost in the noise. At the same time, it can be hard to weigh and assess different information sources based on a Twitter feed.

The third, and again closely related, challenge relates to what we can only describe as the ambiguous status of Citizen Science. The fact that the citizen data scientists talked at times openly about a mix of anxiety, competition, curiosity and a desire for influence tells us that Citizen Science cannot be put in just one epistemological box but will spill over in different ways. Certainly, CDSs report employing data analysis and sharing as a coping or anxiety-reducing device. In wider debate, this has been discussed as one of the additional complications of COVID-19 as loneliness and general anxiety grew during the pandemic (Clothworthy et al., 2021; Dosemagen et al., 2022; Petersen et al., 2021; Varga et al., 2021). The combination of the private and the public, data and discussion, policy and practice might pose challenges in terms of how CDS is incorporated with more traditional policy advice. Nevertheless, here we should simply note that 'conventional' science advice is not immune from similar influences and pressures (van der Sluijs, 2024).

Our fourth observation is that CDS relies on people's own skills (or confidence in their own skills) and time availability, and probably therefore leads to a biased selection of perspectives and questions. As one prominent example, we identified few women and no ethnic minorities, although they might have been the groups most affected by the pandemic (Alon et al., 2020; Bapuji et al., 2020; Tai et al., 2021).

Finally, one obvious question concerns the quality of the information generated by this group of citizen scientists. 'Quality' in this context

is of course open to multiple interpretations and perspectives (Irwin, 2019). Citizen Data Science might not offer a perfect representation of public assessments and forms of expertise. However, and as we have suggested above, this is not to deny the considerable value that this form of Citizen Science can possess. Certainly, and on the basis of the case considered above, it is hard to argue that decision-making in a time of crisis would be improved by excluding the perspectives and forms of evidence presented by these CDSs.

Conclusion

Kinchy (2024) has noted both the manner in which the voices of citizen scientists can be marginalised by established institutions and narrowed so as to operate only within a prescribed framework. We share this concern, but have set out to explore the value of Citizen Science to public policy-making largely from the perspective of the citizen data scientists themselves. To that purpose, we have suggested a number of forms that this contribution has taken in our chosen example: from improving data quality and presentation to questioning underlying policy assumptions, and from serving as a barometer of public concern to improving science communication regarding an issue of pressing national interest. Inevitably in a case of this kind, citizen scientists are not a cross-section of the population in question and, as such, they are likely to reflect certain patterns of interest, perspective and social status (here, a preponderance of professional males). However, that in itself does not undermine the value of their contribution – even if it is certainly a point to be kept in mind when considering CDS as a form of public engagement. It is legitimate to ask who is included but also excluded from Citizen Science activities.

Turning our original question around, it seems relevant to inquire as to the kinds of policy process which might best be able to maximise the policy potential of Citizen Science in cases such as this. On the basis of our discussion here, three factors come to mind.

The first is the capacity of policymakers to avoid insulating themselves from public discussion (including the voices of citizen scientists) but to engage with larger audiences – not only in the interests of democracy but also better policy-making.

The second enabling factor involves recognition of perhaps the most fundamental aspect of Citizen Science, namely that public groups can bring knowledge and expertise to key questions as well as their democratic influence.

The third enabling factor takes us back to a point made early in this paper. Just as with other policy inputs, including those from qualified experts, Citizen Science is likely to be an imperfect source, offering both important strengths but also limitations. Rather than simply celebrating or critiquing its contribution, it is timely to evaluate the contribution of the citizen sciences in more contextual, contingent and multi-dimensional terms.

Acknowledgements

We thank Professor Karen Risager for her comments on an early version of this paper. Helpful feedback was received at a seminar hosted by the Centre for Science and Policy at Cambridge University. We are grateful to our anonymous reviewers for their constructive commentaries.

References

Alon T, Doepke M, Olmstead-Rumsey J and Tertil M (2020) T*he impact of COVID-19 on gender equality (No. w26947)*. National Bureau of Economic Research.

Bapuji H, de Bakker FGA, Brown JA, Higgins C, Rehbein K and Spicer A (2020) Business and Society Research in Times of the Corona Crisis. *Business and Society* 59(6): 1067–1078.

Birkin LJ, Vasileiou E and Stagg HR (2021) Citizen science in the time of COVID-19. *Thorax* 76(7): 636–637.

Boisson M (2023) Un public choisi: la constitution d'un forum hybride dans la mise en qualité des données du Covid. Presentation to seminar on 'Les publics de le mise en nombre'. Centre Émile Durkheim, IFRIS, SFdS, October 20th.

Bonney R (1996) Citizen science: A lab tradition. *Living Bird* 15(4): 7-15.

Bowser A, Cooper C, De Sherbinin A, et al. (2020) Still in need of norms: The state of the data in citizen science. *Citizen Science: Theory and Practice*, 5(1): 1–16.

Christensen J G, Askim J, Gyrd-Hansen D, Madsen H B, Østergaard L J and Mortensen P B (2021) *Håndteringen af covid-19 i foråret 2020: Rapport afgivet af den af Folketingets Udvalg for Forretningsordenen nedsatte udredningsgruppe vedr. håndteringen af covid-19*. Folketinget. Available at: https://www.ft.dk/-/media/sites/ft/pdf/publikationer/haandtering-af-covid19-foraar-2020.ashx (accessed November 26, 2024).

Christensen T, Jensen M D, Kluth M, et al. (2023) The Nordic governments' responses to the Covid-19 pandemic: A comparative study of variation in governance arrangements and regulatory instruments. *Regulation & Governance* 17(3): 658-676.

Clotworthy A, Dissing A S, Nguyen T L, et al. (2021) 'Standing together – at a distance': Documenting changes in mental-health indicators in Denmark during the COVID-19 pandemic. *Scandinavian Journal of Public Health* 49(1): 79–87.

Cohen K and Doubleday R (eds) (2021) *Future Directions for Citizen Science and Public Policy*. Cambridge: Centre for Science and Policy.

COVID-19 Map - Johns Hopkins Coronavirus Resource Center (n.d.). Available at: https://coronavirus.jhu.edu/ (accessed November 18, 2022).

Dosemagen S, Kimura AH, Frickel S and Parker A (2022) Disaster, Participatory Science, and Infrastructure. *Citizen Science: Theory and Practice* 7: 14.

Evans R (2022) SAGE advice and political decisionmaking: 'following the science' in times of epistemic uncertainty. *Social Studies of Science* 52(1): 53-78.

Forthun I, Madsen C, Emilsson L et al. (2024) Excess mortality in Denmark, Finland, Norway and Sweden during the COVID-19 pandemic 2020–2022. *European Journal of Public Health* 34(4): 737-743.

Funtowicz SO and Ravetz JR (1993) Science for the post-normal age. *Futures* 25(7): 739-755.

Grey F (2009) The age of citizen cyberscience, *CERN Courier*, 29th April 2009. Available at: https://cerncourier.com/a/viewpoint-the-age-of-citizen-cyberscience/ (accessed April 12, 2025).

Griffith G J, Morris T T, Tudball M J, et al. (2020) Collider bias undermines our understanding of COVID-19 disease risk and severity. *Nature Communications* 11(1): 1-12.

Groom Q, Strubbe D, Adriaens T, et al. (2019) Empowering Citizens to Inform Decision-Making as a Way Forward to Support Invasive Alien Species Policy. *Citizen Science: Theory and Practice* 4: 33.

Gundelund C and Skov C (2021) Changes in angler demography and angling patterns during the COVID-19 lockdown in spring 2020 measured through a citizen science platform. *Marine Policy* 131(May): 104602.

Hajer M (1995) *The Politics of Environmental Discourse: ecological modernization and the policy process*. Oxford: Clarendon.

Haklay M (2013) Citizen Science and Volunteered Geographic Information: Overview and Typology of Participation. In: Sui D, Elwood S, Goodchild M (eds) *Crowdsourcing Geographic Knowledge*. Springer, Dordrecht, pp.105-122.

Haklay M (2021) Why is it so difficult to integrate citizen science into practice? In: Cohen K and Doubleday R (eds) *Future Directions for Citizen Science and Public Future Directions for Citizen Science and Public*. Cambridge: Centre for Science and Policy, pp.109-116.

Haklay M (2022) Mutual Learning Exercise on Citizen Science initiatives – policy and practice. First topic report: Introduction and Overview of Citizen Science. Directorate-General for Research & Innovation, Brussels.

Hecker S, Haklay M, Bowser A, Makuch Z, Vogel J and Bonn A (eds) (2018) *Citizen Science: Innovation in Open Science, Society and Policy*. London: UCL Press.

Hecker S, Wicke N, Haklay M and Bonn A (2019) How Does Policy Conceptualise Citizen Science? A Qualitative Content Analysis of International Policy Documents. *Citizen Science: Theory and Practice* 4: 32.

Hilgartner S (2024) Pandemics and the politics of knowledge. In: Felt U and Irwin A (eds) *The Elgar Encyclopedia of Science & Technology Studies*, Cheltenham: Edward Elgar, pp. 569-578.

Horst M and Irwin A (2010) Nations at Ease with Radical Knowledge: on consensus, consensusing and false consensusness. *Social Studies of Science* 40(1): 105-126.

Irwin A (2018) Citizen science comes of age. *Nature* 562(7728): 480-482.

Irwin A (1995) *Citizen Science: a study of people, expertise and sustainable development*. London and New York: Routledge.

Irwin A (2006) The politics of talk: coming to terms with the "new" scientific governance. *Social Studies of Science* 36(2): 299-320.

Irwin A (2008) STS perspectives on scientific governance'. In: Hackett EJ, Amsterdamska O, Lynch M and Wajcman J (eds) *The Handbook of Science and Technology Studies Third Edition* Cambridge: MIT Press, pp. 583-607.

Irwin A (2015) Citizen science and scientific citizenship: same words, different meanings? In: Schiele B, Le Marec J and Baranger P (eds) *Science Communication Today – 2015: current strategies and means of action*. Nancy: Presses Universitaires de Nancy – Éditions Universitaires de Lorraine, pp.29-38.

Irwin A (2019) Re-making 'quality' within the social sciences: the debate over rigour and relevance in the modern business school.' *The Sociological Review* 67(1): 194-209.

Irwin A, Jensen T E, Jones, KE (2013) The good, the bad and the perfect: criticising engagement practice. *Social Studies of Science* 43(1): 119-136.

Jasanoff S (1990) *The Fifth Branch: science advisers as policymakers*. Cambridge: Harvard University Press.

Jasanoff S, Hilgartner S, Hurlbut J B, Özgöde O and Rayzberg M (2021) *Comparative COVID Response: Crisis, Knowledge, Politics*. Comparative COVID Response Project, Interim, 12 January. Available at: https://compcore.cornell.edu/ (accessed April 12, 2025).

Kasperowski D and Kullenberg C (2019) The many modes of citizen science. *Science & Technology Studies* 34(2): 2-7.

Kimura AH and Kinchy A (2016) Citizen Science: Probing the Virtues and Contexts of Participatory Research. *Engaging Science, Technology, and Society* 2*:* 331–361.

Kinchy A (2024) Citizen science. In: Felt U and Irwin A (eds) *The Elgar Encyclopedia of Science & Technology Studies*. Cheltenham: Edward Elgar, pp. 215-224.

Kvale S (1994) *Interviews: An introduction to qualitative research interviewing*. Thousand Oaks: Sage.

Liberatore A, Bowkett E, MacLeod CJ, Spurr E and Longnecker N (2018) Social Media as a Platform for a Citizen Science Community of Practice. *Citizen Science: Theory and Practice* 3(1): 3, pp. 1–14.

Miles MB, Huberman AM and Saldaña J (2014) *Qualitative data analysis: A methods sourcebook*. 3rd edition. London: Sage.

Murray B, Kerfoot E, Chen L, et al. (2021) Accessible data curation and analytics for international-scale citizen science datasets. *Scientific Data* 8(1): 1–17.

Nielsen J H and Lindvall J (2021) Trust in government in Sweden and Denmark during the COVID-19 epidemic. *West European Politics* 44(5–6): 1180–1204.

Our World in Data. (n.d.). Available at https://ourworldindata.org/ (accessed November 18, 2022).

Ottinger G (2021) Becoming Infrastructure: Integrating Citizen Science into Disaster Response and Prevention. *Citizen Science: Theory and Practice* 7(1): 1–13.

Petersen MB, Christiansen LE, Bor A, et al. (2021) Communicate Hope to Motivate Action Against Highly Infectious SARS-CoV-2 Variants. *PsyArXiv* 1(1): 1–4. https://psyarxiv.com/gxcyn/

Pueyo T (2020) Coronavirus: The Hammer and the Dance - What the Next 18 Months Can Look Like, if Leaders Buy Us Time. *Medium.* Available at: https://tomaspueyo.medium.com/coronavirus-the-hammer-and-the-dance-be9337092b56 (accessed January 20, 2022).

Roberts S (2020) Flattening the Coronavirus Curve. *The New York Times*. Available at: https://www.nytimes.com/article/flatten-curve-coronavirus.html (accessed January 20, 2022).

Roman D, Reeves N, Gonzalez E, et al. (2021) An analysis of pollution Citizen Science projects from the perspective of Data Science and Open Science. *Data Technologies and Applications* 55(5): 622–642.

SAPEA, Science Advice for Policy by European Academies (2019) *Making sense of science for policy under conditions of complexity and uncertainty*. Berlin: SAPEA.

Sleigh J, Amann J, Schneider M and Vayena E (2021) Qualitative analysis of visual risk communication on twitter during the Covid-19 pandemic. *BMC Public Health* 21: 1-12.

SSI-Dashboard: SSI's smitte- og vaccinedashboard (2021) Available at:https://experience.arcgis.com/template/099eb5c9acea4e18b411997815be2f98 (accessed January 5, 2024).

Sprague D and Tory M (2012) Exploring how and why people use visualisations in casual contexts: Modelling user goals and regulated motivations. *Information Visualisation* 11(2): 106–123.

Tai DBG, Shah A, Doubeni CA, Sia IG and Wieland ML (2021) The disproportionate impact of COVID-19 on racial and ethnic minorities in the United States. *Clinical Infectious Diseases* 72(4): 703-706.

Trajkova M, Alhakamy A, Cafaro F, Vedak S, Mallappa R and Kankara SR (2020) Exploring casual COVID-19 data visualisations on Twitter: Topics and challenges. *Informatics* 7(3): 1–22.

UKEOF Citizen Science Working Group (2021) *Impacts of COVID-19 on UK environmental citizen science Lessons learned and recommendations for the future*.

Van der Sluijs J (2024) Science advice. In: Felt U and Irwin A (eds) *The Elgar Encyclopedia of Science & Technology Studies*, Cheltenham, UK and Northampton: Edward Elgar, pp. 397-406.

Van Oudheusden M (2021) Citizen science in pandemic times: Lessons across east and west. In: Cohen K and Doubleday R (eds) (2021) *Future Directions for Citizen Science and Public Policy*. Cambridge: Centre for Science and Policy.

Van Oudheusden M and Abe Y (2021) Beyond the grassroots: Two trajectories of "citizen scientization" in environmental governance. *Citizen Science: Theory and Practice* 6(1): 1-15.

Van Oudheusden M, Suman AB, Huyse T, Huyse H and Medvecky F (2024) The valuable plurality of the citizen sciences. *Science & Technology Studies* 37(1): 10-20.

Varga TV, Bu F, Dissing AS, et al. (2021) Loneliness, worries, anxiety, and precautionary behaviours in response to the COVID-19 pandemic: a longitudinal analysis of 200,000 Western and Northern Europeans. *The Lancet Regional Health - Europe 000*. 100020.

Webster A (2007) Crossing boundaries social science in the policy room. *Science, Technology, & Human Values* 32(4): 458-478.

Zambach S (2021) Citizen Produced Visualisations as Qualifying the Public Debate on the COVID Pandemic. Poster session presented at *ACM Collective Intelligence Conference 2021*, Frederiksberg, Denmark.