This is an author accepted manuscript version of: Rempel, E. S., Barnett, J. and **Durrant, H**. (2018). Public engagement with UK government data science: Propositions from a literature review of public engagement on new technologies, *Government Information Quarterly*, 35 (4), pp. 569–578. https://doi.org/10.1016/j.giq.2018.08.002

Public engagement with UK government data science: Propositions from a literature review of public engagement on new technologies

HIGHLIGHTS

We develop five propositions for public engagement on government data science.

1. Consider the 'publics' who may be engaged in government data science.
2. Do not assume providing information will lead to public acceptance.
3. Determine contingencies of trust
4. Design deliberative engagements.
5. Ensure holistic public participation that moves beyond privacy and consent.

1. INTRODUCTION

While data is nothing new, in recent years there has been a revolution in the mechanisms of creating, collecting, processing, connecting, and applying data. In the United Kingdom (UK), the recent passing of the Investigatory Powers Act (Investigatory Powers Act 2016, 2016; MacAskill, 2016) evidences that government, in particular, is attentive to these new possibilities that data affords. Indeed, it has a large stake in the development and use of 'data science'. Traditionally government is positioned as a regulator of new technologies such as genetic modification and nanotechnology (M Kearnes, Macnaghten, & Wilsdon, 2006; J Stilgoe, 2007), however it is both the producer and consumer of data science in the UK. This metaphorically muddies the waters since data science both informs government policy yet must also be regulated by government. Whilst private sector corporations tend to use data science to expand their business interests (N.A., 2016), governments use data relating to citizen activity to inform decision making across domains ranging from the collection of household waste to tracking terrorist activity (MacAskill, 2016).

A range of recent high profile events such as Care.Data (Carter, Laurie, & Dixon-Woods, 2015) have shown that government data use can be judged as unacceptable, and in the case of Edward Snowden reveal unprecedented invasions of privacy and civil rights (Greenwald, MacAskill, & Poitras, 2013). This also demonstrates how individuals are uniquely tied to data science as a technology due to them often providing the data that government use. Building on this, the argument that this paper advances is that the precarious basis for government self-regulation heightens the importance of including diverse public voices in the realm of data science. This paper contributes five propositions developed from a thematic synthesis of public engagement literature on science and technology for how and why government should include these public voices. The focus of these propositions is on government activities, however there are of course many ways in which publics self-organise

and themselves seek conversations on data science with government (Braun & Konninger, 2017; Lezaun & Soneryd, 2007; Selin et al., 2017). While we do not offer propositions for these publics, we recognise that the high profile events mentioned above are often driven by publics and open spaces for these same publics to take part in government activity.

In the light of previous cases of public disaffection around the introduction of new technology, as well as early examples of publicity around data being shared in ways that breach public expectations, it is unsurprising that government has started to initiate public engagement processes around data science. We seek to inform our consideration of the potential for public engagement with data science through a critical analysis of engagement and trust building in other new technologies. In the background, we first establish definitional clarity on public engagement and data science before outlining the interface between data science and the citizenry. We then consider some early examples of public disquiet around government data practices. Following this we describe the methodology used in the narrative literature review. In the first half of the discussion, we outline five themes drawn from literature around new technology and public engagement. The second half of our discussion develops these themes into five propositions for data science public engagement in government. We situate the second half of the discussion within examples of how government has begun to engage around data science to consider the relevance of previous public engagement practice to government data science, as well as how we may use the tenets of decentred governance to stimulate theoretical reflection on these propositions.

2. BACKGROUND INFORMATION

2.1 Definitions and conceptions of public engagement

Broadly engagement is defined as the inclusion of publics in some aspect of the development or regulation of policy or technology (Rowe & Frewer, 2005). Rowe and

Frewer (2005) categorise public engagement as communication, consultation, or participation. In these three categories the nature of public involvement changes from informing publics about some aspect of a technology to the participation of publics in the deliberation and development of technology. In an alternative classification pertaining more to the motivations of the instigator of engagement than to the methods used to engage, Fiorino (1990) and later Stirling (2004) describe engagement as either normative, substantive, or instrumental. Under a normative perspective publics are seen as having a democratic right to engagement in relation to a new technology, for example by virtue of living in an area affected by its siting. A substantive perspective positions publics as able to make a difference by virtue of their engagement. For example, publics could be involved in developing a new transport system and provide insight about features that would be most valued. Adopting an instrumental perspective would include publics as a means to an end (Barnett, Cooper, & Senior, 2007), for example as a mechanism of minimising controversy or unwanted media attention.

While these are useful frameworks, there is no prominent theory of public engagement and thus engagement encapsulates a broad range of public-technology interactions. We define public engagement as a subset of democratic activity that focuses specifically on the inclusion of non-technical publics in the development and governance of new technologies. In the case of data science in the UK, this includes a range of activities related to including publics in the development and regulation of government data usage that are not government officials, data scientists, nor other technical publics.

2.2 Defining data science

While traditionally defined as the processes of combining and applying data, we use the term data science to represent *both* the substance and application of data. This includes

both novel types of data like big data and novel applications of data. Included within data science is the concept of open data which refers to data sets that are released in the public domain (Hand, 2013). Definitions of big data are more contentious with early classifications relating to particular characteristics, famously first as the three Vs of velocity, volume, and variety (McNeely & Hamh, 2014) but it has also been defined according to its political and social implications by Markus and Topi (2015, p. 3) who call for "a sociotechnical perspective, viewing Big Data as a cluster or assemblage of data-related ideas, resources, and practices." We consider data science to include both the technical practices of data usage and data technology development, but also the ways that data science interacts with and informs social and political practice. Data science is a term typically used in the UK, thus the literature we draw on and the public engagement propositions developed are most relevant to the UK context. We include some international literature, mostly from the US, but the arguments made here are based on the UK context. The UK has a rich history of public engagement, as well as a strong and burgeoning interest in data-based decision-making. Thus these dual interests in public engagement and data science provide a highly relevant context for considering public engagement on data practices. It is vital to be clear that our focus is on how publics may engage with the regulation and potential applications of data science and not public engagement using data. We are not focussing on how data technologies can be used as a *means* of engagement, for example government using social media for communication or citizens using social media to organise protests (Lee & Kwak, 2012; Warren, Sulaiman, & Jaafar, 2014), but rather how public engagement on data science and its various contingencies can be done.

2.3 Governments, citizens, and the future of data science

From Twitter feeds and sentiment analysis, to store card data and marketing, data and data science are claimed to hold vast potential to improve the efficiency of government

processes and enhance policy development (Gov.UK Blog, 2015). Complex analytics are being used to predict likely crime scenes in Los Angeles while Facebook 'likes' have been piloted for public health monitoring in Florida (Gittelman et al., 2015; PredPol, 2015). The UK Department for Education recently proposed a register in England which would link diverse data on every school and student (Vale, 2016) while, the Investigatory Powers Act (2016) allows for full-scale web-scraping of all citizens' Internet use (MacAskill, 2016). There is not only a proliferation of data but of applications through which government can surveil and ultimately regulate the citizenry. And as highlighted through the Snowden case, these applications are not always deemed to be in the public interest.

The routines, interactions, and practices of citizens are inevitably intertwined – albeit often unknowingly - with the production and use of data science. As a citizenry our actions are translated in to data in a multitude of ways (Johnson, 2014). For example, by directly filling out administrative forms and clicking links online but also through entering a car park that uses ground sensors. Governance and policy-making for health, employment, banking, transport, education, justice, and housing, to various degrees, involve the production of data that can in theory contribute to better outcomes for citizens and government (Joseph & Johnson, 2013). Such outcomes might include enhancing how we interact with various organisations which can in turn realise benefits of saving time or money (Hancock, 2016). However, our relationships with data science may be less passive, for example utilising online information about housing to decide where to live to interactive mapping exercises like Map Kibera that take advantage of local knowledge to overcome shortfalls in government transparency (Donovan, 2012). In this sense, data science, government, and the citizenry exist in a complex data ecosystem with varying, and perhaps decreasing in the case of the citizenry, levels of power and influence. A key message from these tightly intertwined processes is that publics have a fundamental stake in the results and development of data

science techniques, whether that is in the improvement of that technology or the regulation of its use.

2.4 Early evidence of public disquiet around government use of data

In common with the introduction and development of a range of technologies, public disquiet and disaffection with the use of public data by governments has coalesced around several high profile incidents in recent years. In late April 2016, New Scientist reported that the Google-owned machine learning firm DeepMind had unprecedented access to all patient records from London hospitals run by the NHS Royal Free Trust (N.A., 2016) resulting in concern with regard to the unconstrained sharing of private medical records. In 2014, following concerns around the sharing of personal medical information, the NHS data strategy Care.Data was cancelled (Boseley, 2016). Publics are often uncomfortable with commercial access to private data, and prefer data science that has a clear element of public good (Cameron, Pope, Clemence, & Ipsos MORI Social Research Institute, 2014; Davidson et al., 2013; Ipsos MORI, 2016). This suggests interest from publics in the kinds of data that government collect and how they are ultimately used. Hence, data science has entered in to conversations around public engagement with data science acceptability, privacy, and consent (Gov.UK Blog, 2015; Sciencewise, 2012, 2014).

In light of these early examples of public concern, as well as the precarious basis for government self-regulation of data science, we argue for a more nuanced study of public engagement in data science. We advocate, of course, for transparency on what government does with data but the answer to the relatively simple question of whether individuals should know what is happening to their data is assumed to be yes. Publics should know what happens to their information, particularly when data science based technologies have the potential to do harm. Due to these potential social and societal effects, data is a sociotechnical

object and there are critical questions around the interactions between society, government, and data science that need to be queried and deliberated by publics. Thus we now move on first to present our methodology and second to discuss how the history of public engagement with new technologies can help to develop models for public engagement with government data science.

3. METHODOLOGY

Our aim is to examine the potential for public engagement with data science through a narrative literature review of public engagement with other technologies. A narrative literature review focuses on critically developing new models by synthesising the 'most significant items in the field' (Grant & Booth, 2009, p. 94), and does not formally assess the quality of the literature included. A narrative review was chosen due to the broad nature of the research question (e.g. any new technology and public engagement of any form) as well as the desire to include texts that were not empirical in the review. Our main focus is on synthesising lessons from these texts to develop the propositions for government data science public engagement rather than to systematically analyse the literature. While the public engagements synthesised in this paper cover a range of different technologies, these technologies have all faced the same question now confronting government data science. Namely, how to effectively include publics in the development and governance of sociotechnical objects? We take this indirect approach as data science is still a new technology and subsequently there is not a large literature on public engagements with data science as of yet.

Table 1: Search terms and databases

Search terms used	
Public	Public* OR People* OR Societ* OR Communit* OR Populace OR Citizen* OR Person* OR Patient*
Engagement	Engagement OR Consultation OR Dialogue OR Involvement OR Discourse OR Participation OR Communicaton[1]
	Risk* OR Perception* OR Acceptance* OR View* OR Opinion* OR Knowledge OR Attitude* OR Awareness OR Impression* OR Viewpoint*
Trust	Trust* OR Confidence
New Technology	"Emerging *Technolog*" OR "New *Technolog*" OR "Modern *Technolog*" OR "Novel *Technolog*" OR "Developing *Technolog*" OR "Rising *Technolog*"
Databases searched	Scopus, Web of Science, and PsycINFO

1 This word was spelled incorrectly in the original database search.

We extracted relevant literature based on a review of the abstract and title, including both peer-reviewed and grey literature. Search terms used and databases searched are presented in Table 1. Relevance was decided based on whether the article focused on any kind of public opinion-seeking or engagement with a new technology. The technology had to involve some kind of science application. We also included literature suggested by the second author based on her expert opinion. Trust was used due to expert advice and it's prevalence in early reading around nanotechnology engagement, see Wynne (2006) and Walls et al. (2004). The articles were read for themes common to engagement, we followed Braun and Clarke's (2006) methods of reviewing, coding, categorising, and re-categorising key themes. We extracted common themes from a total of 49 articles. A full list of these 49 articles is available in Appendix 1. Articles were included until thematic saturation was reached, i.e. that no new themes were found in new articles. Primary themes were first developed by the lead author then grouped and revaluated by the team. An initial 12 themes were identified which were then combined and simplified which are presented in Table 2. These themes were

then used to develop five propositions for public engagement with government data science. These propositions and their connecting theme are presented in Table 3. We further consider the value of the propositions by then evaluating their relevance to public engagement and governance theory. While Rowe and Frewer (2005) and Fiorino (1990) provide useful frameworks of public engagement, they do not provide a theoretical foundation for thinking about how public engagement, as a concept, fits in government and the governance of technology. In this vein, Braun and Konninger (2017) call for public engagement to be evaluated through a holistic, theoretical lens of decentred governance of science and technology. Engagement is then not about finite, defined public consultations but is underlined by the view of politics "as activities of struggle and conflict concerning the meaning of particular issues that inevitably involve power relations" (pg. 10). They suggest "to take controversy and contestation as a point of departure and study how controversies, publics and issues are brought into being" (pg. 11). Considering the propositions identified in the literature, we evaluate how they can be linked to key tenets and principles of decentred governance. The four tenets we draw on are the recognition of power relations, allowances for pluralism, moving the regulation of technology outside of central government structures, and examining sites of contestation as sites of public engagement (Griggs, Norval, & Wagenaar, 2014).

Table 2: Initial and final thematic analysis of public engagement literature results

Initial twelve themes	1) Trust versus robustness as the purpose of engagement. 2) Public understanding leads to success of new technologies. 3) Critiques of the view that people and trust halt technological development. 4) Transparency will lead to trust. 5) Deliberation is conceived as small group discussions. 6) Mistrust in technology is actually distrust in government.

	7) New technologies will have unique public risks, unlike other technologies thus context is important. 8) Stereotypes of the evil market versus the irrational activities in engagement activities. 9) Public involvement (participation) will lead to success of new technologies. 10) Risk is multifaceted. 11) Upstream deliberation is key to public engagement but is often not done in practice. 12) Publics and other stakeholders often have unshared socio-technical imaginaries of the future.
Five final themes	1) Conceptions of the public in engagement. 2) A knowledge deficit and other early science-citizen interactions. 3) Trust and trustworthiness in public engagement. 4) How aims predict methodologies in public engagement. 5) Imaginaries of the future in public engagement

4. A REVIEW OF PUBLIC ENGAGEMENT WITH NEW TECHNOLOGY

This first half of our discussion introduces the five common themes found in the literature on public engagement with new technology, while in the second half we develop these themes into propositions for government-driven public engagement with data science. Examples of data science engagement and applications are therefore also presented in Section 5. We include literature around various new technologies that are often applications of scientific principles. In line with Nightingale (2014) in considering technologies, we include both the infrastructure and context around technologies that are in essence part of the technologies themselves. For additional reading on themes in public engagement see Braun and Konninger (2017) and Smallman (2016). First, we discuss conceptions of publics that may be involved in engagement. Second, we outline early initiatives of science-society interactions including the development of the Deficit Model. Next, we discuss lessons for engagement in terms of trust building. Our fourth section describes various methodologies of

engagement and how we might move beyond event-based engagement. Finally, we discuss imaginaries of the future and how these impact the potential for public engagement.

4.1 How to define the 'public'

Public engagement, necessarily, involves explicitly or implicitly defining who the public are by deciding who to engage with. Engagement is often about "finding ways of connecting with people who could be mobilised as supporters" (Walker, Cass, Burningham, & Barnett, 2010, p. 942). Various positions are taken relating to whether publics are finite pre-existing groups or constructed and dynamic. Renn (2008) argues for four categories of publics: stakeholders, affected publics, observing publics, and the general public where the general public is the whole or unaffected public. In contrast, Newman (2011) suggests that publics do not pre-exist, rather that public leaders call upon or create publics for a given purpose. Within engagements, the 'real' public is often assumed to be the supportive, silent majority (Burningham, Barnett, & Walker, 2014). In the case of renewable energy technology, Burningham et al. (2014) found publics who oppose new technology were framed as an irrational minority. For example, a common narrative places activists as an unreasonable leftist minority and demonises industry as representing rightist market forces (Barnett, Burningham, Walker, & Cass, 2010; Laurent, 2007; Shelley-Egan & Davies, 2013; Torgersen & Schmidt, 2013). Laurent (2007) similarly found these tendencies to resort to stereotypes were a common feature of public-industrial interactions around nanotechnology. Government, and other technology stakeholders, often have imaginations of who the public are and what they may feel prior to starting any form of engagement. Furthermore, the mental models that are held of publics and the attributions made about their interests, capabilities, and likely behaviours will determine the nature and extent of the engagement opportunities that are provided (Barnett et al., 2010). Thus, defining 'the public' also defines the spaces for public engagement.

4.2 A knowledge deficit and other early science-citizen interactions

The early history of public-technology interactions, later termed public engagement, were framed around the deficit model. This model posits that citizens have a deficit of knowledge around science and technology, and that lack of knowledge drives rejection of technology and scientific ideals. Sturgis and Allum (2004, p. 56) suggest "[implicit] in this programmatic agenda is the claim that 'to know science is to love it'." The deficit model emerged from early science-citizen interactions around technologies such as nuclear power and bioactive substances like thalidomide (Wynne, 2006). In the case of thalidomide, public concerns arose from associations of a causative link between a particular drug and birth defects with more general concerns around science and technology. However, in relation to technologies such as nuclear power, there was an underlying assumption that concern arose from a lack of understanding and knowledge (Grove-White et al., 2004; Wynne, 2006). The "assumption was that no rational and properly informed person could possibly disagree with the desirability of whatever science endorsed – nuclear power, chemical pesticides, chlorofluorocarbons" (Wynne, 2006, p. 215). Following this logic, the method to improve public acceptance (and to ensure the unimpeded progress of such technologies) is to educate and inform the public. This synthesis of communication (Rowe & Frewer, 2005) with instrumental engagement (Fiorino, 1990) was seen in initiatives to inform the general public about technologies to thus ensure their acceptance.

While Sturgis and Allum (2004, p. 55) point out that a "scientifically literate citizenry is also one that can effectively participate in public debates about science", critics of the deficit model take issue with the belief that knowledge about technology guarantees acceptance (M Kearnes et al., 2006; J Stilgoe, 2007; J Stilgoe, Irwin, & Jones, 2006; Wilsdon, Wynne, & Stilgoe, 2005; Wynne, 2006). As Williams, Macnaghten, Davies, and Curtis (2015, pp. 98-99) argue in their recent critique of deficit-like assumptions in fracking

engagement, there "is no guarantee that more information will lead to greater acceptance, or that the availability of facts will lead to a more 'rational' and calculative form of choice-making." In fact, the deficit model reduces publics to mere receivers of information. As Wynne (2006, p. 215) emphatically characterizes, the deficit model is "dogmatically authoritarian and arrogantly self-centred." Knowledge provision, while a component of public engagement, is not sufficient to engage publics in any significant way.

4.3 Trust and trustworthiness in public engagement

In the wake of, and as a response to the dominance of the deficit model, engagement exercises were seen as a method of building trust in new technologies (Burningham et al., 2014; Grove-White, Macnaghten, & Wynne, 2000; Groves, 2011; M. Kearnes & Wynne, 2007; Marris & Rose, 2010; J Stilgoe, 2007; The Global Environmental Change Programme, 1999; Torgersen & Schmidt, 2013; Wilsdon et al., 2005). Trust, or a "firm belief in the reliability, truth, or ability of someone or something" (Oxford University Press, 2017), is often cited as key to successful science-citizen interactions. As Warburton (2009, p. 32) reflected after the failure of a nuclear power consultation, an "atmosphere of hostility, caution and anxiety is not conducive to the flexible and creative environment that is ideal for the design and delivery of engagement activities." A lack of trust is problematic. The role of trust is reflected in the genetically modified organism (GMO) and bovine spongiform encephalopathy (BSE) events of the mid-1990s where publics' concern arose around the potentially harmful effects of genetically modifying crops and poor food safety. Trust was diminished following minimal communication and transparency on uncertainties in these technologies, prompting what the UK House of Lords (2000) deemed a crisis of trust in science. By the mid-2000s, engagement around trust building for things like nanotechnology was commonplace (Groves, 2011; M Kearnes et al., 2006).

Critics of trust building argue it is premised on the same faulty assumption as the deficit model, i.e. the public are a problem (M. Kearnes & Wynne, 2007; J. Stilgoe, Lock, & Wilsdon, 2014). As Groves (2011, p. 792) describes, trust building "sees technology and expertise as threatened by dynamics of distrust that disrupt what is imagined as a pre-existing condition of consensual trust in the promises of science". The assumption is that "the prevailing deficit of public confidence…will be improved through the direct consultation and active engagement [of] lay-public concerns" (M. Kearnes & Wynne, 2007, p. 133). Public involvement then becomes about filling a deficit of trust (Burgess, 2014; Groves, 2011; Involve, 2015; Ipsos MORI, 2006; M. Kearnes & Wynne, 2007). A focus on trust can, therefore, be seen as merely another instrumental mechanism of preventing public controversy (Krütli, Stauffacher, Flüeler, & Scholz, 2010; Marris & Rose, 2010). As J Stilgoe et al. (2006, p. 20) argue, "[the] focus on trust turns the problem into one of communication. And the deliberate attempt to manufacture trust can look deeply untrustworthy." Public engagements that focus on building trust risk doing the opposite. While trust is a component of public and government interactions, it should not be the aim of public engagement. Similar to communication, trust is necessary but not sufficient.

J Stilgoe et al. (2006, p. 21) advocate that "[we] must instead focus on what goes into building trustworthiness". This move from trust to trustworthiness was central to shifting public engagement away from the 'public as burden' argument. Instead of the onus being on publics to gain trust, regulators and developers are instead responsible for inspiring trust. There are several key features of trustworthiness, a common argument being that transparency will lead to trust (Flynn, Ricci, & Bellaby, 2012; M. Kearnes & Wynne, 2007; O'Hara, 2012; Shelley-Egan & Davies, 2013; Stebbing, 2009). Transparency is public openness in the development and regulation of new technologies, but also in the engagement process itself, e.g. trust that public voices will in fact be heard and not marginalised during

public engagements (Warburton, 2009). This openness is often operationalised as releasing data sets, reporting of minutes, and various other practices of being as clear as possible in how something is developed and conclusions about it are drawn. Stebbing (2009, p. 41), in her review of the potential for nanotechnology engagement, argued for "governance that is based on transparency and accountability at the local rather than global level". Thus the concepts of transparency and accountability are closely linked. The argument is if governance and development are open then government and industry are accountable for good practice in technology development. While transparency and openness in innovation are components of good practice, particularly in government, transparency rests on the assumption that if the processes behind innovation are more widely publicised, developers *might* do the right thing (M. Kearnes & Wynne, 2007). Transparency is not a replacement for public engagement, it merely facilitates the potential for successful engagement to take place.

A second feature of being trustworthy is understanding that trust is multifaceted. For example, Walls, Pidgeon, Weyman, and Horlick-Jones (2004) suggest that trust in government agencies is rarely simple, it is a critical trust. Publics have neither blind belief nor total scepticism in government agencies. Instead "perceptions and understandings of government agencies and departments are vague and…susceptible to contingencies of events" (Walls et al., 2004, p. 145). For example, Barnett et al. (2007) found that having a stronger belief in public efficacy was associated with having greater trust in genetic science but conversely lower trust in government. Trust becomes attached to different events, histories, and organisations. Walls et al. (2004, p. 135) call for "a situation whereby trust is provisionally conceptualised as multi-faceted, potentially dynamic, and dependent upon a range of contextual variables." Trustworthy practice requires consideration of these different histories and contexts of individuals and publics who may interact with new technologies, as well as the complexity of public views on government.

4.4 Linking the aims and methodologies of public engagement

If trust and communication are only components of engagement, and not the goal, then what is the aim of public engagement? Groves (2011, p. 787) suggests building robustness, i.e. "produc[ing] forms of technology which are explicitly embedded in different ways of domesticating uncertainty". A focus on robustness aims to encourage public participation that emphasises negotiating technological development and governance and creates feedback mechanisms between stakeholders and publics. This shifts engagement away from 'public as problem' and towards 'public as partner' (J Stilgoe, 2007). Robustness aligns, to some degree, with Fiorino's (1990) substantive and normative aims of engagement. Normative engagement positions publics as due a role in technological development while substantive engagement posits that publics can improve technology. Hence, publics can provide a unique perspective on the purposes and design of technologies and thus build social and technological robustness.

The aim of engagement is fundamentally linked to how to engage. For communication or trust building purposes, engagement could be typified by providing information. With a normative aim, engagement could include public workshops and early consultation on public opinion and under a substantive viewpoint engagement could involve working groups that ask publics to help develop and design a new technology. Building robustness requires both substantive and normative methodologies of engagement. Bonney, Phillips, Ballard, and Enck (2016) highlight the importance of offering multiple sites for participation, e.g. offering more than one way and one time period for people to engage. They suggest that publics are empowered through involvement in question development thus so-called upstream deliberation is key to public engagement (Burri & Bellucci, 2008; Grove-White et al., 2004; Jones et al., 2014; M Kearnes et al., 2006; Pidgeon, Harthorn, Bryant, & Rogers-Hayden, 2009; The Global Environmental Change Programme, 1999; Willis &

Wilsdon, n.d.; Wilsdon et al., 2005). Burri and Bellucci (2008, p. 387) highlight the "consensus that the public should be involved in deliberative discussions and assessments of emerging technologies at a much earlier stage of technological developments." However, authors caution that engagement should be at the right time rather than simply early as early engagements can result in "participants [that] are difficult to find and to engage, they often discuss half-heartedly in an artificial setting" (Torgersen & Schmidt, 2013, p. 52) In other words, if individuals are engaged under a normative or substantive aim then that engagement must have a meaningful impact on technological development.

The vast majority of engagement exercises reviewed in developing this paper were time-limited small group discussions, such as workshops or focus groups (Burri & Bellucci, 2008; Flynn et al., 2012; Ipsos MORI, 2006; Jones et al., 2014; M Kearnes et al., 2006; Pidgeon et al., 2009; J Stilgoe, 2007; Walker et al., 2010). Pidgeon et al. (2009) and M Kearnes et al. (2006), for example, conducted workshops that included presentation, discussion, and hypothetical deliberation around nanotechnology. Flynn et al. (2012) held Citizens' Panels around hydrogen energy that used presentation, discussion with experts, and again hypothetical deliberation. In this literature, public engagement is often one-off events. That is not to say that all engagement is still event-based, drawing from policy literature that is not on the topic of science and technology but nevertheless involves some form of government-public interaction, forms of digital engagement are emerging (Gagliardi et al., 2017; Panagiotopoulos, Al-Debei, Fitzgerald, & Elliman, 2012). An example is Lauriault and Mooney's (2014) descriptions of crowdsourcing mapping and other forms of collaborative work. A clear lesson is that these kinds of more involved, participatory engagements are preferable over the older one-off event-style engagements, and that both kinds of engagement persist in today's government-citizen interactions.

4.5 Imaginaries of the future for data science engagement

Engagement generally falls around two main topics: regulation and development. The former being more common than the latter. For example, Flynn et al. (2012) asked participants to debate the potential for a hydrogen economy and drew out themes on hazards, risks, trust, and regulation. Similarly, Burri and Bellucci (2008, p. 388) asked participants to discuss the potential for nanotechnology with an aim at stimulating debate and "to help decision makers in assessing nanotechnologies." While development is a rarer topic, it can be framed by theories of co-design and participatory research (Cornwall & Jewkes, 1995), where in essence publics are engaged through the technology design process itself. Burgess (2014) describes a hybrid where publics were engaged in discussions around genetic technologies and involved in governance along the research process. However, what is missing in these topics is the very first stage of development, i.e. what technologies are developed and why?

Groves (2011), M Kearnes et al. (2006) and J Stilgoe et al. (2006) describe these visions of the future as sociotechnical imaginaries, or narratives of how society envisions the future. These imaginaries create the boundaries for what is and is not acceptable in technological development, and are developed through complex socio-technical relationships between industry, academia, media, publics, and government. Groves (2011) argues that the dominant imaginary is an *empty future horizon*. This empty future is one where anything is possible and if anything is possible then there are no limits on development or developer autonomy. However, if there are no limits then what role is there for publics to shape what the future looks like? Some authors argue that ambivalence around new technology often masks a deeply engrained fatalism in public engagement exercises (Grove-White, Macnaghten, Mayer, & Wynne, 1997; M. Kearnes & Wynne, 2007). Essentially publics do not feel like they can make a difference. Groves (2011, p. 792) goes as far as to call upon industry and government to conceive of engagement "as a constitutive part of a democratic rewriting of the contract between strategic techno-science and society." The challenge is to

allow engagement to be about negotiation and not consultation or communication. Only by first allowing publics to have input on what the future can look like, can more specific engagement, such as regulation or development occur.

5. PROPOSITIONS FOR GOVERNMENT-DRIVEN DATA SCIENCE

In the second half of our discussion, we address how each of the themes identified in the literature can be used to build a single proposition for government data science public engagement. In doing so we reflect on current and previous government-based data science engagements as well as how the theory of decentred governance can be applied to better understand the theoretical dimensions of the propositions. The connection between the literature review themes and the propositions is presented in Table 3 where each proposition was developed from one theme found in the literature. While, the previous section focussed on discussing and qualifying the themes we synthesised, this next section turns to evaluating their relevance to data science in government. We do so by drawing on examples of data practice in government as well as a key underpinning theory relevant to including publics in governance, i.e. decentred governance.

Table 3: Literature review lessons and propositions for government data science in the UK

Public Engagement Literature Theme Identified		**Proposition for Government Data Science**
Variations in how to define the 'public'	→	Consider the varied and many 'publics' who may be engaged in government data science.
The focus on the knowledge deficit model in early science-citizen interactions	→	Do not assume providing publics with information on data science initiatives will lead to public acceptance.
A shift towards a focus on trust and trustworthiness in public engagement	→	Determine the contingencies of trust for government data science and public engagement through trustworthy practice
How the aim impacts the methodology of public engagement	→	Design public engagements that incorporate robust, critical, and ongoing deliberation of data science

The importance in considering imaginaries of the future for data science engagement	→	Ensure holistic public participation that moves beyond privacy and consent

5.1 Proposition One: Consider the varied and many 'publics' who may be engaged in government data science

While identifying the potential publics in any data science project ultimately rests on the specifics of the project itself, it is nevertheless important to consider the potential publics that one may engage with. As with any engagement exercise there are both vast and discrete potential publics, however in practice it is possible to find a subset of the public with whom it is more relevant to engage. Therefore, policy-makers initiating public engagement will benefit from reflecting on how to best characterise and access relevant publics in line with their specific purpose. Data science can also be used to segment populations to identify previously uncategorised groups, for example children in need of educational assistance or individuals most likely to default on a loan (Ginnis et al., 2016; Joseph & Johnson, 2013). And while there are positive and negative aspects to identifying and grouping individuals in this way, it nevertheless requires critical reflection on how to ensure these new publics are both aware of and participants in data science. There are increasingly publics that are already organised and self-identified in relation to data science. Civic hacking groups, in essence public professionals, and associated hackathons offer, in theory, a resource for citizens with non-technical skills to engage with data science. For example, a local civic data group in the UK, Bath: Hacked, recently involved a group of 23 volunteers in a public exercise to create online accessibility maps. Similarly, Lauriault and Mooney (2014) describe a range of group mapping exercises from crowdsourcing to more intensive citizen science. Considering the many and varied publics allows for a range of different opinions to come forward on data science, this is closely linked to the decentred governance tenet of pluralism. Pluralism

doesn't require a consensus from these groups, rather it allows for the reality of the many different ways of looking at an issue. Government must, therefore, include a range of publics in data science engagement to fully capture the pluralism of data science governance.

5.2 Proposition Two: Do not assume providing publics with information on data science initiatives will lead to public acceptance

Events related to Care.Data illustrate that discussions around government data science were predicated on a deficit model. Care.Data was a data sharing programme that proposed linking general practice records across England. In early 2013, NHS England ran a leaflet information campaign to inform the public of the Care.Data initiative. The NHS assumed that providing knowledge would be sufficient to establish, as Carter et al. (2015) describe, a social license to reuse medical data. However, the poor quality of the information provided, lack of public consultation, and unclear opt-out mechanisms led to public and media concern and many patients requesting to opt out of any sharing of their medical records (Carter et al., 2015; Kirby, 2014). Subsequently, the entire programme was abandoned (Boseley, 2016). Care.Data demonstrates the faulty assumptions around knowledge provision and public acceptance, as well as providing a cautionary tale around government self-regulation. Further to this highly publicised case, a public and professional consultation on data sharing conducted by Cameron et al. (2014, p. 5) suggested that "many participants were sceptical of the value of informing the general public about the ADRN [administrative data linkage] initiative" citing that participants felt publics "would not understand such a complex topic through simple messages, and thus would become worried about data security and privacy when there is not necessarily a reason to be." The notion that even communication is not needed in data initiatives is concerning. We link these deficit-like assumptions to power relations in decentred governance. Assuming that informing publics is enough to gain acceptance, is if nothing else, a reflection on the relative power between government, data

scientists, and publics where publics are not provided with the option to critique. No matter how laudable data sharing initiatives may be, government must allow for spaces of public participation that recognise these power imbalances. This also demonstrates a form of public engagement where publics self-organised around an issue to contest government practice. This is both an example that exposes the imbalance of power between government and publics while also demonstrating publics reasserting their call for power in data science governance (Braun & Konninger, 2017). And while there is recognition in the literature that citizens need decision-making power in data-facilitated engagements like e-participation platforms, (Attard, Orlandi, Scerri, & Auer, 2015; Gagliardi et al., 2017), in practice engagement on data governance is still limited to communication.

5.3 Proposition Three: Determine the contingencies of trust for government data science and public engagement through trustworthy practice

Debates around trust, accountability, and transparency have also begun in the realm of data science. One key form of new data, open data, rests on the ideal of transparency (Levy & Johns, 2016; O'Hara, 2012; Schrock, 2016). While open data covers any form of publicly accessible dataset, O'Hara (2012, p. 4) argues that it specifically offers "the possibility of holding government accountable." A key focus in the UK government's recently announced algorithmic decision-making inquiry is "how algorithmic decision-making can be conducted in a 'transparent' or 'accountable' way" (Commons Select Committee, 2017). Transparency in data practices is connected to ideals of accountability while also enabling people to engage with data science (Attard et al., 2015). However, this assumes publics have the resources, skills, and finances to use and interpret the data provided. Arguably, with highly specialist technologies like data science, this is often not the case (Levy & Johns, 2016). In practice, as Attard et al. (2015, p. 414) reflects, while "the benefits of open data outweigh the efforts required, it appears there is a lack of public participation in open government initiatives." The

open data to accountability perspective has largely been critiqued as being naïve, in fact Johnson (2014) argues that 'opening up data can function as a tool of disciplinary power' (pg. 270) and further suggests open data theorists move to a perspective of *information justice*, which 'can be especially useful in overcoming the capabilities gaps between enterprises and individuals…and make information pluralism a reality (p. 272). While not an argument for public engagement per se, it is a perspective that allows for the socio- to reintegrate to technical systems.

In terms of considering the various contingencies and contexts of trust in data science, it is challenging to divorce views of government from views of government data practice. In a public consultation on data linking, Cameron et al. (2014, p. 14) found that "participants trusted government's intentions more than commercial companies" and yet "[participants] were also worried about personal data being leaked, lost, shared or sold by government departments to third parties." They concluded that "[l]ow trust in government more generally seemed to be driving these views" (Cameron et al., 2014, p. 22). It is challenging to distinguish what publics may use as reference points to evaluate data risks. Cameron et al. (2014) suggest that media events, like the revelations of Edward Snowden and Julian Assange, drive the belief that data can never be truly secure. In a widely publicised case in the UK, Google subsidiary DeepMind developed a collaboration agreement with the NHS Royal Free Trust gaining access to millions of identifiable patient records. While they claimed to only be accessing these records to provide an app that identifies acute kidney (AKI) patients, as Powles and Hodson (2017) discuss there were in fact no real limitations on what they could do with the data, nor any transparency in what they were actually doing. In addition, patients were enrolled based on a principle of implied consent due to the app being involved in direct care, beyond being a shaky assumption on any grounds, implied consent would only be in effect for those patients with AKI (Powles & Hodson, 2017). After the New

Scientist revealed the problematic latitude of this agreement (N.A., 2016) Google and DeepMind published press releases highlighting the small scope of the project in order to reassure publics. Despite, or perhaps due to, these minimal *post hoc* attempts at public participation this data science collaboration remains under scrutiny.

It is events like these, where untrustworthy practice comes to light, that are likely to drive public perceptions of data science. UK governments have made some efforts to define how publics feel about data. Examples include the Cabinet Office's Government Digital Service consultation on what the 'red lines' are in ethical use of data science (Gov.UK Blog, 2015; Sciencewise, 2012, 2014). In 2015, the Office for National Statistics and the Wellcome Trust held a series of deliberative workshops with general publics and specialist groups to identify how people felt about commercial access to health data (Ipsos MORI, 2016). Government-commissioned reports on public views of data have been implemented or funded by the Economic and Social Research Council , Ipsos MORI , the Scottish Government, Sciencewise , the Cabinet Office and more (Cameron et al., 2014; Davidson et al., 2013; Gov.UK Blog, 2015; Ipsos MORI, 2006, 2016; Sciencewise, 2012, 2014). While these activities represent a step toward establishing publics' views on various aspects of data science, they do not equate nor reflect public influence. Despite how they may be designed around finding the ethical lines of data science, these are not sufficient for public engagement nor do they overcome untrustworthy practice in other areas. In fact these consultations evidence a lack of any kind of decentred governance. Particularly in the Google DeepMind case, they evidence attempts to 'get around' public engagement rather than critical inward reflections on trustworthy practices that could enable public engagement. Thus the main tenet of decentred governance that regulation of data science would occur beyond central government can only be built upon trustworthy practice.

5.4 Proposition Four: Design public engagements that incorporate robust, critical, and ongoing deliberation of data science

As data is already intertwined with publics, creating more robust models of public ownership of personal data could move governments beyond simply conducting one-off events. The ongoing embedded nature of public data practices calls for designing innovative engagement exercises that reflect how data science is increasingly a part of day-to-day life. Opportunities for involving publics in data collection, use, and governance abound. Examples include citizen science initiatives that crowd-source data processing (Bonney et al., 2016; Lauriault & Mooney, 2014) and the aforementioned hackathons that leverage the knowledge of 'public professionals' in typically government-led events to drive data innovation (Sousa, 2013). Historical exercises in the US like citizen engagement through commenting and rule-making could, of course, extend to data policies and practices. These efforts allow for open public comments on policy drafts, see Johns and Saltane (2016). However, there is the potential for more inclusive forms of engagement with data than these limited one-off events. Innovative forms of engagement using social media are already in practice (Lee & Kwak, 2012), however it is important to note that these are often engagements using data technologies and not about data science itself.

There is also an important lesson in thinking critically about the aim of these data science engagements. Substantive forms of engagement should have the possibility to make a difference and to enable publics to have impact on the processes of data science. As Malik (2013, p. 6) describes, "[the] first step of the journey toward Big Data governance involves stakeholder engagement". As government is facing increasingly complex challenges in how to regulate their own use of data science technologies, innovative and multiple venues for public participation can help government address broader questions around ethical and beneficial data science. The technical knowledge required for conducting data science is

obviously also a limitation, and thus focussing on ethical questions would allow broader engagement. This will require new and innovative forms of public engagement that allow for sites of contestation and pluralism, this 'means that participation in science governance takes place in a multiplicity of sites' (Braun & Konninger, 2017, p. 10). And while it is impossible to predict where these sites of contestation will develop, public engagement that is robust, critical, and ongoing allows public engagement to be less about shutting down public opposition and more about opening up debate (Stirling, 2004).

5.5 Proposition Five: Ensure holistic public participation that moves beyond privacy and consent

What is the future imaginary of data? Thus far the topics of data science engagement have focussed on privacy and consent (Joseph & Johnson, 2013; Schintler & Kulkarni, 2014; Stough & McBride, 2014). For example, Cameron et al.'s (2014) Dialogue on Data sought "to explore public understanding and views of administrative data and data linking". Limiting public engagement to discussions around privacy and consent sets a tight boundary for public influence, and can be seen as a way of shutting down potential areas for conversation and contestation. We suggest a future horizons perspective where there is an opportunity for government to involve publics in creating an imaginary of the data future. More than upstream deliberation or building robustness, government can think about the way publics can be involved in the ideology that drives the use of data science in policy, and the kinds of data that are created. This can be achieved through more inclusive and early engagement. As Kennedy, Poell, and van Dijck (2015, p. 6) argue, '[to] participate in datafied social, political, cultural and civic life, ordinary people need to understand what happens to their data, the consequences of data analysis, and the ways in which data-driven operations affect us all.' We would go further to add that 'ordinary people' need to be understood as the key stakeholder in a datafied world. Publics need to be engaged in how they imagine data, what

sorts of information they see as useful, how they think data could be used, and most importantly in how they wish the future to look. Only through this kind of decentralized governance with data science can the issue of government self-regulation be addressed.

6. CONCLUSION

6.1 Strengths, limitations, and academic contribution

This is the first article to offer propositions for government data science public engagement practice that are rooted in concrete and empirical lessons from public engagement literature. This is also the first narrative literature review of public engagement with new technology to critically describe and synthesise these lessons. We suggest both academic scholars and government could build on these lessons to develop effective public engagement activities. A potential limitation of this approach is that we did not conduct a systematic review of the literature and the lessons found here could be strengthened and enhanced by further systematic reviews of public engagement in the future. As well this article focuses on introducing the reader to the field of public engagement, a more critical and theory-driven approach could be used in future. Particularly, an approach that reviewed public-government interactions in areas not related to science and technology may be useful. And finally, while this article focussed on engagement related to data science, these lessons may be useful to other fields of data and digital technologies including concepts like big data, open data, and machine learning.

6.2 Final thoughts

The increasing use of data science in policymaking is creating new spaces for public engagement. These new opportunities can create confusion on how and where to effectively engage publics in the development and regulation of data science. We synthesised five themes from 49 articles that focussed on public engagement with new technology. These

themes were then used to develop five novel propositions for public engagement with government data science. This includes considering the varied and many 'publics' who may be engaged in government data science, not assuming that providing publics with information on data science initiatives will lead to public acceptance, determining the contingencies of trust for government data science and public engagement through trustworthy practice, and designing public engagements that incorporate robust, critical, and ongoing deliberation of data science. Our final proposition is to ensure holistic public participation that moves beyond privacy and consent. This highlights the importance of recognising that publics have an interest in *how* and *why* government uses data science. In particular they are due a role in decided what government *should* use data science for. Government has a unique opportunity to allow publics in decision-making spaces around how data is created, collected, and utilised for the good of society. Data is a public matter. It is the next steps that government take that will decide whether publics are adversaries or partners in this data future.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing Interests

None Declared.

References

Attard, J., Orlandi, F., Scerri, S., & Auer, S. (2015). A systematic review of open government data initiatives. *Government Information Quarterly, 32*(4), 399-418. doi:10.1016/j.giq.2015.07.006

Barnett, J., Burningham, K., Walker, G., & Cass, N. (2010). Imagined publics and engagement around renewable energy technologies in the UK. *Public Understanding of Science, 21*(1), 36-50. doi:10.1177/0963662510365663

Barnett, J., Cooper, H., & Senior, V. (2007). Belief in public efficacy, trust, and attitudes toward modern genetic science. *Risk Analysis, 27*(4), 921-933. doi:10.1111/j.1539-6924.2007.00932.x

Bonney, R., Phillips, T. B., Ballard, H. L., & Enck, J. W. (2016). Can citizen science enhance public understanding of science? *Public Underst Sci, 25*(1), 2-16. doi:10.1177/0963662515607406

Boseley, S. (2016). NHS to scrap single database of patients' medical details. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2016/jul/06/nhs-to-scrap-single-database-of-patients-medical-details

Braun, K., & Konninger, S. (2017). From experiments to ecosystems? Reviewing public participation, scientific governance and the systemic turn. *Public Underst Sci*, 963662517717375. doi:10.1177/0963662517717375

Burgess, M. M. (2014). From 'trust us' to participatory governance: Deliberative publics and science policy. *Public Underst Sci, 23*(1), 48-52. doi:10.1177/0963662512472160

Burningham, K., Barnett, J., & Walker, G. (2014). An Array of Deficits: Unpacking NIMBY Discourses in Wind Energy Developers' Conceptualizations of Their Local Opponents. *Society & Natural Resources, 28*(3), 246-260. doi:10.1080/08941920.2014.933923

Burri, R. V., & Bellucci, S. (2008). Public perception of nanotechnology. *Journal of Nanoparticle Research, 10*(3), 387-391. doi:10.1007/s11051-007-9286-7

Cameron, D., Pope, S., Clemence, M., & Ipsos MORI Social Research Institute. (2014). *Dialogue on Data - Exploring the public's views on using linked administrative data for research purposes*. Retrieved from London:

Carter, P., Laurie, G. T., & Dixon-Woods, M. (2015). The social licence for research: why care.data ran into trouble. *J Med Ethics, 41*(5), 404-409. doi:10.1136/medethics-2014-102374

Commons Select Committee. (2017). Algorithms in decision-making inquiry launched. Retrieved from https://www.parliament.uk/business/committees/committees-a-z/commons-select/science-and-technology-committee/news-parliament-2015/algorithms-in-decision-making-inquiry-launch-16-17/

Cornwall, A., & Jewkes, R. (1995). What is participatory research? *Social Science and Medicine, 41*(12), 1667-1676.

Davidson, S., McLean, C., Treanor, S., Ipsos MORI, Aitken, M., Cunningham-Burley, S., . . . University of Edinburgh. (2013). *Public acceptability of data sharing between the public, private and third sectors for research purposes*. Retrieved from

Donovan, K. (2012). Seeing like a slum: Towards open, deliberative development. *Georgetown Journal of International Affairs, 13*(1), 97-104.

Fiorino, D. J. (1990). Citizen participation and environmental risk: A survey of institutional mechanisms. *Science Technology Human Values, 15*(2), 226-243.

Flynn, R., Ricci, M., & Bellaby, P. (2012). Ambiguity, complexity and uncertainty surrounding the hazards of hydrogen and public views of emergent risks. *Journal of Risk Research, 15*(4), 373-387. doi:10.1080/13669877.2011.634517

Gagliardi, D., Schina, L., Sarcinella, M. L., Mangialardi, G., Niglia, F., & Corallo, A. (2017). Information and communication technologies and public participation: interactive maps and value added

for citizens. *Government Information Quarterly, 34*(1), 153-166. doi:10.1016/j.giq.2016.09.002

Ginnis, S., Evans, H., Boal, N., Davies, E., Palmqvist Aslaksen, A., Ipsos MORI Social Research Institute, . . . Sciencewise. (2016). *Public dialogue on the ethics of data science in government*. Retrieved from London:

Gittelman, S., Lange, V., Gotway Crawford, C. A., Okoro, C. A., Lieb, E., Dhingra, S. S., & Trimarchi, E. (2015). A new source of data for public health surveillance: Facebook likes. *J Med Internet Res, 17*(4), e98. doi:10.2196/jmir.3970

Gov.UK Blog. (2015). Data Science ethics. Retrieved from https://data.blog.gov.uk/2015/12/08/data-science-ethics/

Grant, M. J., & Booth, A. (2009). A typology of reviews: an analysis of 14 review types and associated methodologies. *Health Info Libr J, 26*(2), 91-108. doi:10.1111/j.1471-1842.2009.00848.x

Greenwald, G., MacAskill, E., & Poitras, K. (2013). Edward Snowden: the whistleblower behind the NSA surveillance revelation. *The Guardian*.

Griggs, S., Norval, A. J., & Wagenaar, H. (2014). Introduction : Democracy, conflict and participation in decentred governance. In S. Griggs, A. J. Norval, & H. Wagenaar (Eds.), *Practices of Freedom: Decentred Governance, Conflict and Democratic Participation*. New York, NY: Cambridge University Press.

Grove-White, R., Kearnes, M., Miller, P., Macnaghten, P., Wilsdon, J., & Wynne, B. (2004). *Bio-to-Nano? Learning the lessons, interrogating the comparison (amended version)*. Retrieved from London:

Grove-White, R., Macnaghten, P., Mayer, S., & Wynne, B. (1997). *Uncertain world: Genetically modified organisms, food and public attitudes in Britain*. Retrieved from Lancaster:

Grove-White, R., Macnaghten, P., & Wynne, B. (2000). *Wising up: The public and new technologies*. Retrieved from Lancaster:

Groves, C. (2011). Public engagement and nanotechnology in the UK: restoring trust or building robustness? *Science and Public Policy, 38*(10), 783-793. doi:10.3152/030234211x13070021633440

Hancock, M. (2016). *Data science ethical framework launch: Matt Hancock speech*. London: Cabinet Office Retrieved from https://www.gov.uk/government/speeches/data-science-ethical-framework-launch-matt-hancock-speech.

Hand, D. J. (2013). *Data, not dogma: Big data, open data, and the opportunities ahead.* Paper presented at the Advances in Intelligent Data Analysis XII, London UK.

Investigatory Powers Act 2016. (2016). *Investigatory Powers Act 2016*. Retrieved from United Kingdom:

Involve. (2015). *Written evidence submitted by Involve (BIG0010)*. Retrieved from London:

Ipsos MORI. (2006). *The use of personal health information in medical research general public consultation: Final report*. Retrieved from London:

Ipsos MORI. (2016). *The one-way mirror: Public attitudes to commercial access to health data*. Retrieved from London:

Johns, M., & Saltane, V. (2016). *Citizen engagement in rulemaking: Evidence on regulatory practice in 185 countries*. Retrieved from

Johnson, J. A. (2014). From open data to information justice. *Ethics and Information Technology, 16*(4), 263-274. doi:10.1007/s10676-014-9351-8

Jones, A. R., Anderson, A. A., Yeo, S. K., Greenberg, A. E., Brossard, D., & Moore, J. W. (2014). Using a deliberative exercise to foster public engagement in nanotechnology. *Journal of Chemical Education, 91*(2), 179-187. doi:10.1021/ed400517q

Joseph, R. C., & Johnson, N. A. (2013). Big data and transformational government. *IT Professional, 15*(6), 43-48.

Kearnes, M., Macnaghten, P., & Wilsdon, J. (2006). *Governing at the nanoscale: People, policies and emerging technologies*. Retrieved from London:

Kearnes, M., & Wynne, B. (2007). On nanotechnology and ambivalence: The politics of enthusiasm. *NanoEthics, 1*(2), 131-142. doi:10.1007/s11569-007-0014-7

Kennedy, H., Poell, T., & van Dijck, J. (2015). Data and agency. *Big Data & Society, 2*(2). doi:10.1177/2053951715621569

Kirby, T. (2014). Controversy surrounds England's new NHS database. *The Lancet, 383*(9918), 681. doi:10.1016/s0140-6736(14)60230-0

Krütli, P., Stauffacher, M., Flüeler, T., & Scholz, R. W. (2010). Functional-dynamic public participation in technological decision-making: site selection processes of nuclear waste repositories. *Journal of Risk Research, 13*(7), 861-875. doi:10.1080/13669871003703252

Laurent, B. (2007). Diverging convergences - Competing meanings of nanotechnology and converging technologies in a local context. *Innovation-the European Journal of Social Science Research, 20*(4), 343-357. doi:10.1080/13511610701760804

Lauriault, T. P., & Mooney, P. (2014). *Crowdsourcing: A geographic approach to public engagement*. Retrieved from

Lee, G., & Kwak, Y. H. (2012). An Open Government Maturity Model for social media-based public engagement. *Government Information Quarterly, 29*(4), 492-503. doi:10.1016/j.giq.2012.06.001

Levy, K. E. C., & Johns, D. M. (2016). When open data is a Trojan Horse: The weaponization of transparency in science and governance. *Big Data & Society*, 1-6. doi:10.1177/2053951715621568

Lezaun, J., & Soneryd, L. (2007). Consulting citizens: technologies of elicitation and the mobility of publics. *Public Understanding of Science, 16*(3), 279-297. doi:10.1177/0963662507079371

MacAskill, E. (2016). 'Extreme surveillance' becomes UK law with barely a whimper. *The Guardian*. Retrieved from https://www.theguardian.com/world/2016/nov/19/extreme-surveillance-becomes-uk-law-with-barely-a-whimper

Malik, P. (2013). Governing big data: Principles and practices. *IBM Journal of Research and Development, 57*(3/4), 1-13.

Markus, M. L., & Topi, H. (2015). *Big data, big decisions for science, society, and business: Report on a research agenda setting workshop*. Retrieved from Washington, D.C.:

Marris, C., & Rose, N. (2010). Open engagement: Exploring public participation in the biosciences. *PLoS Biol, 8*(11), e1000551. doi:10.1371/journal.pbio.1000551

McNeely, C. L., & Hamh, J. (2014). The Big (Data) Bang: Policy, Prospects, and Challenges. *Review of Policy Research, 31*(4), 304-310.

N.A. (2016). Revealed: Google AI has access to huge haul of NHS patient data. *New Scientist*. Retrieved from https://www.newscientist.com/article/2086454-revealed-google-ai-has-access-to-huge-haul-of-nhs-patient-data/#.VyOORsEjCiU.twitter

Newman, J. (2011). Public leadership as public-making. *Public Money & Management, 31*(5), 315-322. doi:10.1080/09540962.2011.598336

Nightingale, P. (2014). *What is technology? Six definitions and two pathologies*. Retrieved from Brighton:

O'Hara, K. (2012). *Transparency, open data and trust in government: Shaping the infosphere*. Paper presented at the 4th Annual ACM Web Science Conference.

Oxford University Press. (Ed.) (2017) Oxford English Dictionary. Oxford: Oxford University Press,.

Panagiotopoulos, P., Al-Debei, M. M., Fitzgerald, G., & Elliman, T. (2012). A business model perspective for ICTs in public engagement. *Government Information Quarterly, 29*(2), 192-202. doi:10.1016/j.giq.2011.09.011

Pidgeon, N., Harthorn, B. H., Bryant, K., & Rogers-Hayden, T. (2009). Deliberating the risks of nanotechnologies for energy and health applications in the United States and United Kingdom. *Nature Nanotechnology, 4*(2), 95-98. doi:10.1038/nnano.2008.362

Powles, J., & Hodson, H. (2017). Google DeepMind and healthcare in an age of algorithms. *Health and Technology*. doi:10.1007/s12553-017-0179-1

PredPol. (2015). Predictive Policing: The predictive policing company.
Renn, O. (2008). Risk Participation. In O. Renn (Ed.), *Risk Governance: Coping with Uncertainty in a Complex World*. London: Earthscan.
Rowe, G., & Frewer, L. J. (2005). A typology of public engagement mechanisms. *Science, Technology & Human Values, 30*(2), 251-290. doi:10.1177/0162243904271724
Schintler, L. A., & Kulkarni, R. (2014). Big data for policy analysis: The good, the bad, and the ugly. *Review of Policy Research, 31*(4), 343-348.
Schrock, A. R. (2016). Civic hacking as data activism and advocacy: A history from publicity to open government data. *New Media & Society, 18*(4), 581-599.
Sciencewise. (2012). Public dialogue on data openness, data re-use and data management: Sciencewise co-funded project 2012. Retrieved from http://www.sciencewise-erc.org.uk/cms/public-dialogue-on-data-openness-data-re-use-and-data-management/
Sciencewise. (2014). Big Data: Public views on the collection, sharing and use of personal data by government and companies.
Selin, C., Rawlings, K. C., de Ridder-Vignone, K., Sadowski, J., Altamirano Allende, C., Gano, G., . . . Guston, D. H. (2017). Experiments in engagement: Designing public engagement with science and technology for capacity building. *Public Underst Sci, 26*(6), 634-649. doi:10.1177/0963662515620970
Shelley-Egan, C., & Davies, S. R. (2013). Nano-industry operationalizations of "responsibility": Charting diversity in the enactment of responsibility. *Review of Policy Research, 30*(5), 588-604. doi:10.1111/ropr.12032
Smallman, M. (2016). Public understanding of science in turbulent times III: Deficit to dialogue, champions to critics. *Public Understanding of Science, 25*(2), 186-197.
Sousa, S. (2013). *Lessons learnt from a public–private big data hackathon*. Retrieved from London:
Stebbing, M. (2009). Avoiding the trust deficit: Public engagement, values, the precautionary principle and the future of nanotechnology. *Journal of Bioethical Inquiry, 6*(1), 37-48. doi:10.1007/s11673-009-9142-9
Stilgoe, J. (2007). *Nanodialogues: Experiments in public engagement with science*. Retrieved from London:
Stilgoe, J., Irwin, A., & Jones, K. (2006). *The received wisdom: Opening up expert advice*. Retrieved from London:
Stilgoe, J., Lock, S. J., & Wilsdon, J. (2014). Why should we promote public engagement with science? *Public Understanding of Science, 23*(1), 4-15.
Stirling, A. (2004). Opening up or closing down? Analysis, participation and power in the social appraisal of technology. In M. Leach, I. Scoones, & B. Wynne (Eds.), *Science and citizens: Globalization and the challenge of engagement* (pp. 218-231). London: Zed Books.
Stough, R., & McBride, D. (2014). Big data and U.S. public policy. *Review of Policy Research, 31*(4), 339-342.
Sturgis, P., & Allum, N. (2004). Science in Society: Re-Evaluating the Deficit Model of Public Attitudes. *Public Understanding of Science, 13*(1), 55-74. doi:10.1177/0963662504042690
The Global Environmental Change Programme. (1999). *The politics of GM food*. Retrieved from Swindon:
Torgersen, H., & Schmidt, M. (2013). Frames and comparators: How might a debate on synthetic biology evolve? *Futures, 48*, 44-54. doi:10.1016/j.futures.2013.02.002
UK House of Lords. (2000). *Select committee on science and technology: Science and society*. Retrieved from London:
Vale, R. (2016). The register of schools in England goes into alpha. Retrieved from https://data.blog.gov.uk/2016/11/29/the-register-of-schools-in-england-goes-into-alpha/
Walker, G., Cass, N., Burningham, K., & Barnett, J. (2010). Renewable energy and socio technical change: Imagined subjectivities of 'the public' and their implications. *Environment and Planning A, 42*(4), 931-947.

Walls, J., Pidgeon, N., Weyman, A., & Horlick-Jones, T. (2004). Critical trust: understanding lay perceptions of health and safety risk regulation. *Health, Risk & Society, 6*(2), 133-150. doi:10.1080/1369857042000219788

Warburton, D. (2009). *Evaluation of BERR's engagement of the public and other interested parties in the future of civil nuclear power in the UK: Final Report*. Retrieved from London, UK: http://www.sharedpractice.org.uk/Downloads/Nuclear_report.pdf

Warren, A. M., Sulaiman, A., & Jaafar, N. I. (2014). Social media effects on fostering online civic engagement and building citizen trust and trust in institutions. *Government Information Quarterly, 31*(2), 291-301. doi:10.1016/j.giq.2013.11.007

Williams, L., Macnaghten, P., Davies, R., & Curtis, S. (2015). Framing 'fracking': Exploring public perceptions of hydraulic fracturing in the United Kingdom. *Public Underst Sci*. doi:10.1177/0963662515595159

Willis, R., & Wilsdon, J. (n.d.). *From bio to nano and beyond: A progressive agenda for technology, risk and the environment*. Retrieved from London:

Wilsdon, J., Wynne, B., & Stilgoe, J. (2005). *The public value of science: Or how to ensure that science really matters*. Retrieved from London:

Wynne, B. (2006). Public engagement as a means of restoring public trust in science--hitting the notes, but missing the music? *Community Genet, 9*(3), 211-220. doi:10.1159/000092659