

Data For Policy International Conference 2022
Discussion Paper. Submission 3428.

Authors: **Lucille Tetley-Brown, Angela Daly, Esperanza Miyake, Annie Sorbie**

Title: ***Unlocking public value from personal data? Brokering citizen-centred data-use spaces for the private sector – the Scottish Example***

Research / policy question:

How can or should public authorities permit private companies and others to access and use the personal data held by public authorities? The Scottish Government is addressing this question through its '**Unlocking the Value of Data**' (UVOD) Programme and **associated Independent Expert Group (IEG)**. This presentation / paper offers an account of some of the processes undertaken between March and October 2022 within the UVOD Programme, focusing on the work of the Independent Expert Group (IEG), in order to develop a framework for private sector use of personal datasets held by the public sector. This multi-stakeholder group, (of which three of us are members - Daly, Miyake and Sorbie), includes both academics and practitioners, and is working on developing principles for public sector agencies to use in deciding whether – and how - to grant data access to private companies, guided by notions of public trust, public benefit and public interest. In this contribution we briefly set out how this initiative has operated over a 6+ month period, and offer some insights on the navigation of contested notions, such as public benefit and public interest, as well as some of the challenges that can arise in this space.

Research methodology: We employ a reflective practice method (Yanow, 2009; Leering, 2014) to consider the processes this programme of work is undertaking. As authors we combine interdisciplinary academic approaches, with expertise across media and communications, law, critical studies, and sociology methods. We also engage with legal (especially law in context, see Twining 1997) and policy analysis.

Research data: In this extended abstract we reflect on insights from our own experiences in the programme delivery team. As part of the reflective practice method, we set these insights in the context of the data sharing landscape in Scotland, as well as the wider literature around data usage.

1. Introduction

The public sector in many, if not all, countries hold large amounts of data about people - including citizens, migrants and visitors - which interact with public services such as health, education, social security. Accessing this data - much of which will constitute 'personal data' i.e., data about identified or identifiable individuals - may bring various benefits for research, public service delivery and may also result in corporate profits being generated if private companies can access the data. Granting access to this data may also lead to better transparency and checks and balances within institutions of power in

a country. Various governments, including the Scottish Government, have open data policies and initiatives but this mainly relates to the sharing and opening up of non-personal data. Personal data, regulated in many countries by data protection legislation, constitutional / human rights to privacy and other frameworks, poses more complications for 'opening' up access beyond the public sector and even data sharing within the public sector.

In this discussion paper we outline some of the context, issues and challenges we have encountered in the UVOD and IEG work so far and set out the intended next steps for the IEG and UVOD programme. This paper will reflect on some of the processes used in the UVOD work, focussing on the IEG activities. We will focus on how the related concepts of 'the public interest' and 'public benefit' have been taken account of in the work progressed under the auspices of the Programme, and where the term 'value' is predominantly defined through citizen-centred concerns that include issues such as equality, informed consent, agency, and practice, among others. We use these nuanced terms (and, in particular, 'the public interest' and 'public benefit') to also connect to academic and practice-based discussions about if and how public sector (personal) datasets can be brokered to be accessed by the private sector in ways which generate public trust and produce public benefit. We will highlight any lessons from the Scottish case we are involved in which might inform decision-making in this area more generally.

Please note that these are our views as independent academic researchers and do not represent the views of the IEG and/or the Scottish Government.

2. Background and the creation of the Independent Expert Group

The international dialogue about the value of data (including but not limited to 'Big Data') builds from a starting place of "the more data, the better", where value is placed on issues such as large-scale representation and the democratization of data (Mason & Patil, 2015). This position gets caveated with academic and practitioner insights relating to issues including data quality, management, access, and literacy in the application of data to a set purpose.

Much has been written about making data accessible, and where possible fully open, especially when that data is held by the public sector. The concept of 'Big Data' usage within, or based from, government datasets has received much attention over the past decade (see Klievink et al., 2017; Margetts & Sutcliffe, 2013; Mayer-Schönberger & Cukier, 2013). There is ongoing research to understand the attainment of benefits from open government data (e.g., Zuiderwijk et al., 2019). However, there will always be datasets that must remain substantially closed to unbridled application by certain groups, such as general private sector usage for broad commercial gain. This is due to legislative – as well as moral and ethical – demands, particularly with datasets containing personal data. The situation throws up an interesting conundrum about the scope to derive value from this type of data asset. How do we identify the 'value'? Should such value be derived? By whom? What are the parameters of use and the potential for usefulness, serving the public benefit or interest, in the case of closed data?

To help address this, the Scottish Government has set up the UVOD programme with its associated Independent Expert Group (IEG) which comprises a multistakeholder group of experts from the public sector, private sector, third / charitable sector and academia. Three of us are members of this group and, of these three, one (Daly) is the independent chair of the group. (As mentioned above, we reiterate here that the material in this extended abstract constitutes our own views and not those of the IEG or Scottish Government.) The IEG has been meeting during 2022 and is scheduled to complete its work in early 2023. The purpose of the IEG is to formulate research- and evidence-informed recommendations, guidance, and principles for the Scottish Government to implement in Scottish public authorities to help them in decision-making about whether to grant access to the personal data they hold when they receive requests from private sector companies. The Scottish Government plans for the recommendation, guidance, and principles to be implemented in an operationalisation phase once the IEG is disbanded.

3. Scottish context

With the UVOD Programme, the Scottish Government is seeking to design a citizen-centred approach to the opportunities from facilitation of increased use of public sector held (personal) data. A primary concern is establishing robust governance mechanisms to handle the important topics of data quality, management, and skills in the use of the data. These skills may be more than pertain to data science, and it is argued that they include important social (or 'collective' / public) considerations when dealing with data that is generated without the choice or meaningful consent of individuals (such as in the operation of governmental and public service delivery activities), and where the custodianship of that data is philosophically complicated. Whilst access is a fundamental component, to enable the access within parameters deemed socially defensible – that is, trusted by society - is of utmost concern. In the UVOD Programme, the initiators seek to solve the conundrum of trusted use of public sector held personal data by the private sector. The work is being progressed with a sharp focus on being citizen-led and including public engagement as a key component (Scottish Government, 2022).

The existing data policy landscape in Scotland is complex and multifaceted. Scotland is a devolved nation in the UK, with some matters reserved to the UK Government and some matters over which the Scottish Government has devolved competence. An example of the former is data protection, with the UK-wide Data Protection Act 2018 (containing the UK General Data Protection Regulation - which is still in line with EU data protection standards at the time of writing). This is a key piece of legislation which governs both public and private sector processing of personal data throughout the UK, including Scotland. The Scottish Government has various relevant policies, including an Open Data policy (for non-personal data) from 2015, which the Scottish Government plans to update. For some public sector personal data especially in the public healthcare sector (i.e., Scottish NHS), a network of Trusted Research Environments (TREs) / Safe Havens exist, which facilitate researchers from outside the NHS accessing and using this data in accordance with the 'Seven Safes' framework.

The Scottish Government has a series of related policies and work programmes to UVOD, including its Digital Strategy (Scottish Government, 2021a) and AI Strategy (Scottish Government, 2021b), plus various associated activities, all of which are intended to support the Scottish Government in progress for Scotland to be an 'ethical digital nation'. These activities may signal a different approach from the Scottish Government to data and digital markets than the central UK Government: UVOD may be another element to demonstrating this distinctiveness and, in theory, a more citizen-centred approach.

4. Draft principles

One of the IEG outputs is a series of principles to underpin decision-making by data controllers in the Scottish public sector when faced with requests from private companies to access personal data the controllers hold. The IEG formulated the following draft principles between July and August 2022 and made them public for comment and revision. At the time of writing we are still open to receiving feedback on the principles which may be incorporated in their next, and final, iteration. The draft principles are:

1. Public engagement & involvement
- 2. Public interest & public benefit**
3. Do not harm
4. Precaution
5. Transparency
6. Law, ethics & best practice
7. Right to opt out
8. Enabling Conditions
9. Regular Review

During the course the IEG's work so far, it has become clear that a number of conceptual and practical issues have coalesced around draft principle 2, which (currently) states as follows:

'2. Public interest and public benefit

All access to public sector personal datasets must be done in the public interest and must also produce public benefit.'

As such, it is these two notions - of public benefit and of the public interest - that we turn to next. More specifically, we first very briefly survey some existing research around the use of these terms in the context of various data usages, before drawing out three key implications that this may have for the work of the UVOD programme and the IEG.

5. Public benefit, public interest, and potential implications for the UVOD programme and beyond

5.1 Public benefit

A literature review, which was commissioned for the UVOD programme, on public engagement regarding the use of public sector data by or with the private sector over the last 10 years, both in the UK and internationally, sets the position out very clearly with regard to the importance of public benefit in the data context, in that:

“Demonstrable public benefit is the most prevalent consideration that publics have around private sector access to and use of public sector data. While the definition and scope of ‘public benefit’ is open and contested, publics want to see evidence that public benefit of some kind is the primary driver of public sector data access, that it can actually be achieved, and that it outweighs any possible private benefits.” (Erikainen and Cunningham-Burley, 2021)

This relationship – between acceptability and public benefit – is further illustrated by a Wellcome study (Ipsos Mori, 2017) in relation to health data and public attitudes to commercial access. This specifically considered private sector access to data and the conditions under which this may or may not be permissible, and describes how participants applied four key tests when considering the acceptability of data usages (Figure 1).

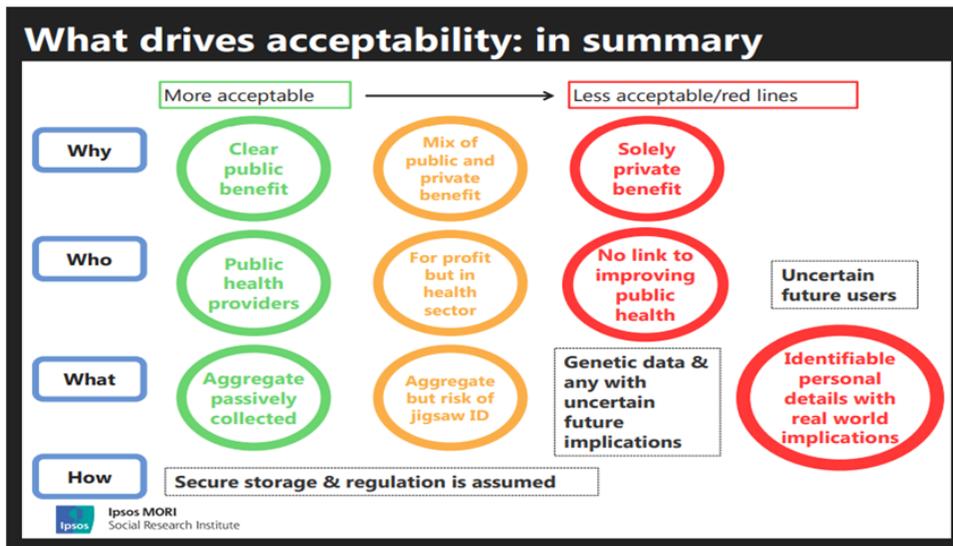


Figure 1 (Ipsos Mori, 2017: 9)

The overarching point that emerges here, for our purposes, is that decisions around acceptability may exist on a sliding scale, with those that have clear public benefit at one end, and those that have solely private benefit at the other. Further, it points to a space where these benefits may be ‘mixed’ in nature.

A recent public dialogue, which was co-funded by the National Data Guardian amongst others, provides a deep dive into public benefit, exploring how this might be assessed in

the data context (Van Mill, 2021). This was conducted in the context of health and care data with around 100 participants, and its findings underline the need for transparency throughout the data lifecycle, and for authentic public engagement with a cross-section of society, amongst other matters.

In the academic sphere, Aitken and colleagues have extensively written about the public engagement work they have conducted in relation to health data sharing, including in the Scottish context (Aitken et al., 2016). In particular, their work in relation to public expectations of public benefits from data-intensive health research (Aitken et al., 2018) has indicated that the term ‘public’ may be construed broadly, so that data usage can benefit as many people as possible. However, understandings of relevant publics may also be needs-led: in other words, there may be broader public benefit in research using data that benefits a smaller group or number of people in need (for example, research in relation to rare diseases). Similarly, Aitken and colleagues found that participants’ preference in terms of the types of benefits was to keep this broad – so, in the context of health research, these benefits were not just seen as medicalisation, but also related to living longer, happier, and healthier lives. Perhaps more notably, publics were also concerned that such benefits should be measurable, and that these would actually be realised through the actions of key policy and government stakeholders.

A strength of the IEG is that its members work across a range of sectors and areas, which has also led to scrutiny of how concepts like public benefit have been addressed elsewhere. For example, official UK Government guidance on the public benefit requirement (Charity Commission, 2013; plus updated format 2017) in the context of Charity Law might bear on understandings of public benefits. This directs us to consider the distinction within the term ‘public benefit’ between its ‘public aspect’ and its ‘benefit aspect’. To satisfy the ‘benefit aspect’ of public benefit “...a purpose must be beneficial” and “...any detriment or harm that results from the purpose must not outweigh the benefit”. To satisfy the ‘public aspect’ of public benefit the purpose must “...benefit the public in general, or a sufficient section of the public” and “...not give rise to more than incidental personal benefit” (Charity Commission 2013: 5). As noted above, this is also a distinction explored in research conducted by Aitken et al. (2018).

5.2 Public interest

To turn next to notions of ‘the public interest’, it is apparent that this term can be equally, if not more, elusive. In the context of health research regulation, it has been claimed that ‘actions taken in the public interest can be broadly described as those that promote objectives valued by society’ (Liminal Spaces, 2021).

More specifically, in the context of data use, the public interest is a prominent feature of the policy and legal regimes that govern the use of confidential data – for example in data protection legislation and the common law duty of confidentiality in the UK. However, neither this legislation nor case law provide a definition of what is, or is not, ‘in the public interest’. Indeed, what emerges from these discussions is that, much like public benefit, the public interest is deeply contextual, and so perhaps we should consider what the public interest ‘does’, rather than solely what it ‘is’, and how it may

relate to other similar terminology, such as the public benefit. In this regard it is of note that the Information Commissioner's Office (ICO), in recent guidance on the research provisions in the UK's DPA 2018 and GDPR has suggested that:

“...the public interest in the context of research should be interpreted broadly to include any clear and positive benefit to the public likely to arise from that research. The public interest covers a wide range of values and principles relating to the public good, or what is in the best interests of society”. (ICO, 2022)

In other research, the connection is made between public interest and public benefit, to argue that a principal function of the public interest is “...to carve out a legally legitimate space within which [research] activities that infringe on individual interests but have potential public benefits can be lawfully conducted, which otherwise would not be permitted” (Sorbie, 2022). However, the argument is also made for a conception of the public interest that is *socially* legitimate, pointing to the difficulties of defining this term on the basis of a homogenised conception of who ‘the public’ are, and in the absence of engagement with actual publics views (for example see Sorbie, 2020; 2021).

5.3 *Potential implications for the UVOD programme*

The whistle-stop tour of the core terminology, as covered in sections 5.1 and 5.2 above, is in no way comprehensive. Nonetheless, there are several important implications for the UVOD programme. Not least, an overarching point that emerges is that a ‘bottom-up’ approach may be preferred, whereby the public interest and public benefit are terms that are co-constructed with publics, in specific contexts, rather than ‘defined’ by the IEG (i.e., a top-down approach). Further, with the aim of co-construction in mind, a number of considerations emerge that might helpfully be explored with publics in the specific context of the UVOD programme.

Let us turn to three prominent considerations.

First, how should we understand ‘publics’, when we use terminology such as ‘public benefit’ and ‘the public interest’ in the context of the UVOD programme? This programme operates, as all governmental initiatives, in circumstances where ‘the public’ is not a homogenous block. As such, this raises questions around how ‘the public’ is understood, for example in broader or narrower terms, as well as directing attention to how unrepresented or underrepresented voices may be captured in circumstances where, as set out in the aforementioned Scottish government commissioned literature review:

“There is no singular ‘public perspective’ on private sector use of public sector data, but rather, while overarching patterns can be identified, publics are plural, and individuals’ views are shaped by a diverse range of intersecting demographic and attitudinal variables.” (Erikainen and Cunningham-Burley, 2021)

Second, how should we understand the ‘interests’ of these publics, and the types of ‘benefits’ that are expected to accrue from the use of public sector data? This may

include consideration of matters such as: the relationship between benefits and harms, and/or the extent to which public and private benefits and interests can and should co-exist.

Third, if it is accepted that these benefits and interests are dynamic rather than static, then a key issue is how these should be understood, evaluated and/or demonstrated over time. This is not only at the point at which decisions about access to data are made prospectively, but also in relation to how these decisions - whether to allow or deny data access - should be reviewed throughout the data use lifecycle.

6. Challenges

In the penultimate section of this discussion paper, we step back and reflect on the discussion above, and consider, in broader terms, three of the challenges that may be relevant to initiatives that engage with the use and non-use of data, such as the UVOD programme.

First, and as demonstrated earlier, polysemous terms such as 'public interest' and 'public benefit' may be difficult to define due to their heterogeneity, because their meanings can differ according to discipline, sector and their relevant specific codes of practice. Beyond issues of nomenclature, such terms are thus driven by economic, socio-cultural and geo-political forces that shape the way they are defined, identified, measured and assessed by different segments of society. The UVOD's IEG is made up of a democratic, multidisciplinary group, approaching issues surrounding personal data and their public use through multiple perspectives. Part of this process is to collectively consider these multiple ways of understanding questions of 'public benefit' and 'public interest'. Yet, the challenge still remains in how best to define these terms that capture their various meanings that reflect the 'on-the-ground reality' in a way that does not value one meaning over the other, whilst also having to produce a set of practical guidelines to encourage the standardisation of ethical data practice and use for all.

At the heart of this conundrum is the very idea of 'value', a key term which is part of the 'Unlocking the Value of Data' title for the Programme itself. How can we begin to define the 'value' of data, when to do so depends on how we understand 'public interest' and 'public benefit' - and when this itself is subject to change? For example, financial value could potentially benefit certain segments of the society; whereas for others, financial value may not only be a low-priority concern, but even the reason for inequality and disparity. In this sense, the difficulty here is not just about grappling with different definitions of value, but also, who has the right to decide what *type* of value should be placed above the other, when considering the 'public benefit' of unlocking 'the value of personal data' for everyone. Whose public interest? What kind of benefit, for whom? How can we even begin to measure and assess value?

The question of multiplicitous values, interests and benefits leads to a related second challenge for the UVOD. Data itself is also complex, and to apply Pasquale's (2015) term, we are largely living in a 'black box society', where data on the whole is neither entirely anonymous nor controllable. Despite data regulations and safeguarding

practices like GDPR in place, data is used and abused daily and in large scale beyond their intended purpose: this is precisely because different segments and sectors of society have different uses and value of data, and the infrastructural boundaries in between are porous and difficult to regulate.

For example, clinically approved digital health tools can potentially be harnessed for public benefit, by unlocking the clinical value of using public, personal data in a trusted research environment: health apps which collected COVID-19 related health data, for example, were used for the containment of the virus as well as to advance clinical knowledge. In this case, it could be argued that the value of unlocking health data was for public interest. Yet, what if these digital tools sit – as they often do – on private data infrastructures? Supportive private companies who are providing data infrastructures may not be collecting personal health data, for example, but they may be collecting other meta-data, such as GPS or internet searches. In such situations, how can we regulate or even monitor how, where and by whom data is used? How can we ensure contextual integrity (Nissenbaum, 2004) remains protected, and data is used ‘correctly’?

Such questions are perhaps impossible to answer or solve, and indeed, are summed up by Pasquale’s own queries: “What are the social implications of the invisible practices that hide the way people and businesses are labelled and treated? How can the law be used to enact the best possible balance between privacy and openness?” (2015: 1) This is the balance which the UVOD needs to consider, and one that needs public engagement and continuous understanding of what is acceptable and what is not acceptable to ‘unlock’, by whom, where and why.

Relatedly - and finally - a third significant challenge which lies ahead – beyond the life and remit of the UVOD – is that in order to address all of the above, we need a continuous dialogue with multiple publics to understand how issues like ‘value’ might change over time. For example, the coronavirus pandemic very quickly changed the different uses, modes of collection and assessment, and ‘values’ attached to public data; what would have been ‘unethical’ before the pandemic, quickly became ‘crucial’ and of public interest and benefit. Similarly, technology and public digital practices – and thus behaviour and perception – change over time. These are things that need continuous evaluation, assessment, public engagement and involvement need a longitudinal approach, not one based on a ‘snapshot’ of society at a given moment in time. It is relevant to consider the mechanisms – including the resources - for such a long-term and ongoing review. In the end, by whom and in what ways is unlocking value of data itself valued?

7. Next steps

The next steps for the IEG are to produce an updated version of the principles with recommendations, guidance, and context in early 2023, before the planned operationalisation stage of UVOD from quarter two of 2023 onwards.

References

- Aitken, M., Cunningham-Burley, S., & Pagliari C. (2016). 'Moving from trust to trustworthiness: Experiences of public engagement in the Scottish Health Informatics Programme'. *Science and Public Policy*, 43(5) 713–723
- Aitken, M., Porteous, C., Creamer, E., & Cunningham-Burley, S. (2018). 'Who benefits and how? Public expectations of public benefits from data-intensive health research'. *Big Data & Society*, 5(2), 2053951718816724.
- Erikainen, S. and Cunningham-Burley, S. (2021) 'Public engagement around the access of public sector data with or by private sector organisations'. Literature Review for Scottish Government, [forthcoming via UVOD webpages]
- Charity Commission (2013) 'Public benefit: the public benefit requirement (PB1)'. [accessed 31-10-22]
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/588234/PB1_The_public_benefit_requirement.pdf
- Information Commissioner's Office, (2022). 'Draft guidance on the research provisions within the UK GDPR and the DPA 2018' [accessed 31-10-22]
<https://ico.org.uk/media/about-the-ico/consultations/4019614/research-provisions-draft-consultation-202202.pdf>
- Klievink, B., Romijn, B. J., Cunningham, S., & de Bruijn, H. (2017). 'Big data in the public sector: Uncertainties and readiness'. *Information systems frontiers*, 19(2), 267-283.
- Leering, M. (2014). 'Conceptualizing reflective practice for legal professionals'. *JL & Soc. Pol'y*, 23, i-106.
- Liminal Spaces, (2021). Concept note 'Public interest in health research' [accessed 31-10-22] <https://www.law.ed.ac.uk/sites/default/files/2021-03/Public%20interest%20concept%20note.pdf>
- Margetts, H., & Sutcliffe, D. (2013). 'Addressing the policy challenges and opportunities of "Big data"'. *Policy & Internet*, 5(2), 139–146. doi:10.1002/1944-2866.POI326.
- Mayer-Schönberger, V., & Cukier, K. (2013). *Big data. A revolution that will transform how we live, work and think*. London: John Murray Publishers.
- Mori, I. (2017). 'The one-way mirror: public attitudes to commercial access to health data.' [accessed 31-10-22]
https://wellcome.figshare.com/articles/journal_contribution/The_One-Way_Mirror_Public_attitudes_to_commercial_access_to_health_data/5616448/1

Lucille Tetley-Brown, Angela Daly, Esperanza Miyake, Annie Sorbie. November 2022.

Nissenbaum, H. (2004). 'Privacy as contextual integrity'. Wash. L. Rev., 79, 119.

Pasquale, F. (2015). *The black box society: The secret algorithms that control money and information*. Harvard University Press.

Patil, D. J., & Mason, H. (2015). *Data Driven*. " O'Reilly Media, Inc."

Scottish Government (2022) 'Unlocking the Value of Data' Programme [accessed 31-10-22]

<https://www.gov.scot/groups/unlocking-the-value-of-public-sector-data-for-public-benefit/>

Scottish Government (2021a) 'A Changing Nation: How Scotland Will Thrive In A Digital World' [accessed 31-10-22] <https://www.gov.scot/publications/a-changing-nation-how-scotland-will-thrive-in-a-digital-world/>

Scottish Government (2021b) 'Scotland's Artificial Intelligence Strategy: Trustworthy, Ethical and Inclusive' [accessed 31-10-22] <https://www.scotlandaistrategy.com/the-strategy>

Sorbie, A. (2022) 'Operationalising 'publicness' in data-intensive health research regulation: An examination of the public interest as a regulatory device' (Doctoral dissertation, University of Edinburgh)

Sorbie, A. (2021). *The Public Interest*. In G. Laurie, E. Dove, A. Ganguli-Mitra, C. McMillan, E. Postan, N. Sethi, et al. (Eds.), *The Cambridge Handbook of Health Research Regulation* (Cambridge Law Handbooks, pp. 65-72). Cambridge: Cambridge University Press. doi:10.1017/9781108620024.009

Sorbie, A. (2020). 'Sharing confidential health data for research purposes in the UK: where are 'publics' in the public interest?'. *Evidence & Policy*, 16(2), 249-265.

Twining, W. (1997) *Law in context: enlarging a discipline*. Oxford University Press.

Van Mil, H. (2021). 'Putting Good into Practice: A Public Dialogue on Making Public Benefit Assessments When Using Health and Care Data.' *National Data Guardian*. [accessed 31-10-22]

<https://www.gov.uk/government/publications/putting-good-into-practice-a-public-dialogue-on-making-public-benefit-assessments-when-using-health-and-care-data>

Yanow, D. (2009) *Ways of Knowing: Passionate Humility and Reflective Practice in Research and Management*. *The American Review of Public Administration*, 39(6), 579-60.

Zuiderwijk, A., Shinde, R., & Janssen, M. (2019). 'Investigating the attainment of open government data objectives: Is there a mismatch between objectives and results?'. *International Review of Administrative Sciences*, 85(4), 645-672.