Centre for Data Ethics and Innovation Consultation Response — defenddigitalme

About defenddigitalme

defenddigitalme is a campaign for children's data privacy and digital rights formed in 2017 in response to concerns from parents and privacy advocates about increasingly invasive uses of children's personal data in education. We support safe, fair and transparent collection and use of pupil data. More information: http://defenddigitalme.com/

---

1. We are responding only to the principle in the consultation executive summary[1] and question 1, Do you agree with the proposed role and objectives for the Centre? under the theme of *"Understanding the public's views, and acting on them, will be at the heart of the Centre's work, as well as responding to and seeking to shape the international debate."*

2. The CDEI must be conscious of its political context, and it is surprising that while it mentions, "*as well as responding to and seeking to shape the international debate*" there is no mention of how it would do this once cut off from European research infrastructure, post-Brexit, and the implications of the potential loss of funding and data adequacy.

3. The answer to the sub-question 8, "*Should the Centre make its activities and recommendations public?*" is a resounding yes.

4. In order to do this and for it to be meaningful, requires a foundational level of public digital understanding[2], and that requires a commitment by government to enable transparency how personal data held in and on behalf of the public sector are used. Without this, the centre will be built on an illegitimate and unstable foundation, without the necessary social license[3] for the promotion of public administrative data exploitation it seeks to deliver.

5. There is a lack of public understanding how personal data as part of public administrative datasets are widely used for indirect purposes by third-parties.

6. The Lords Select Committee report on *AI in the UK*[4] in March 2018, suggested that,*"the Government plans to adopt the Hall-Pesenti Review recommendation that 'data trusts' be established to facilitate the ethical sharing of data between organisations."*

7. The Committee recognised that there is an inherent risk in a lack of public understanding and legitimacy of these plans, because "*under the current proposals, individuals who have their personal data contained within these trusts would have no means by which they could make their views heard, or shape the decisions of these trusts*."

8. *The CDEI* should not operate in a vacuum of what has gone before in public engagement and building UK data infrastructures.

---

[1] Consultation document 'Understanding the public's views' [p5} https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/715760/CDEI_consultation__1_.pdf

[2] Miller C, Coldicutt R and Kitcher H. (2018) People, Power and Technology: The 2018 Digital Attitudes Report. London: Doteveryone. http://attitudes.doteveryone.org.uk/

[3] The social licence for research: why *care.data* ran into trouble Carter, P., Laurie, G., Dixon-Woods, M. (2014) BMJ https://jme.bmj.com/content/early/2015/01/23/medethics-2014-102374

[4] Select Committee on Artificial Intelligence, AI in the UK: ready, willing and Able? (March 2018) https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf

9. The CDEI should be aware of public engagement work carried out in 2013 by Ipsos MORI, as the Administrative Data Research Network set up a new infrastructure for "deidentified" data linkage. Public dialogue[5] was carried across across the UK. Surveys and polls and 14 face-to-face workshops, concluded the same as was apparent to everyone at care.data engagement events in 2014-15[6].

10. *The Public Dialogue on Data* engagement work concluded there is not public support for:

- *"Creating large databases containing many variables/data from a large number of public sector sources,*
- *Establishing greater permanency of datasets,*
- *Allowing administrative data to be linked with business data, or*
- *Linking of passively collected administrative data, in particular geo-location data"*

11. All of the above were seen as having potential privacy implications or allowing the possibility of re-identification of individuals within datasets.

12. *"The other 'red-line' for some participants was allowing researchers for private companies to access data, either to deliver a public service or in order to make profit. Trust in private companies' motivations were low."*[7]

13. All of the above could be central to how the Centre considers "innovation" as a commercial undertaking. An industry driven push to exploit the public's personal data, must not mean yet again, that they are the last to know, and that cannot be legitimised by selected public panels which are not representative and inevitably under pressure to support the aims and work of the Centre.

14. This *Public Dialogue* was a sub-set of the wider *Public Attitudes to Science (PAS) 2014* project[8] undertaken by Ipsos MORI for the then UK Government Department of Business, Innovation and Skills (BIS) and the Economic and Social Research Council (ESRC).

15. The PAS survey questions asked about on numerous potential uses of big data, including operational and research purposes. The PAS findings show that, on balance, the public oppose personal data [in public administrative data] being used for commercial gain.

16. The linking of anonymous government administrative data to better tailor public services garners relatively little support (56%), a*nd concluded, "While a majority seem to be relatively unconcerned about the use of their records in 'big data' analysis, there is strong opposition to some of the specific ways in which private companies might operationalise this data."*

17. Despite this, public services are listed as a target area for AI and the Centre to support.

---

[5] Ipsos MORI commissioned by the ESRC and ONS https://adrn.ac.uk/media/1245/sri-dialogue-on-data-2014.pdf Public Dialogue on Data, 2014, regards the establishing of the ADRN p.59-60

[6] care.data engagement – is it going to jilt citizens after all? A six month summary in twenty-five posts. (2014) http://jenpersson.com/care-data-postings-summary/

[7] Public Dialogue on Data, 2014 Ibid p.59-60

[8] Public Attitudes to Science (PAS) 2014 projechttps://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/348830/bis-14-p111-public-attitudes-to-science-2014-main.pdf

18. Defining "*Accurate targeting of public services to those most in need and more effective distribution of public resources" (1.14)* is deeply subjective. "Accurate" and "effective" by whose measures? These applications of public policy can be deeply unethical and the Centre should be aware of the public reservations around this work.

19. The public good is not a clear cut purpose in itself, and interests change with the politics of the day[9]. A Centre that is used to rubber stamp unethical uses of public data, by the political choices of any given government, will be neither innovative nor ethical.

20. The Lords Select Committee on AI perhaps unintentionally highlighted that data sharing today is often *un*ethical, as it suggested that driving ethical uses of data, would be one aim of Data Trusts. However "*a set of relationships underpinned by a repeatable framework, compliant with parties' obligations'",* seems little better than the status quo, and will cement a focus on data use, not respect for the autonomy, rights and the human dignity of individuals and communities.

21. More people marking their own homework, [if the Centre is steered by data users approving how they distribute data and its uses] will simply bake-in the problems we have today. These could include MOUs and data sharing agreements between government departments which facilitate decidedly *un*ethical policy and practices[10] but are deemed politically to be in the public interest.

22. Such policy choices made by government and public bodies, directly contradict the CDEI consultation statement (1.13), *"We have a strong tradition in data and AI innovation, careful navigation and management of the ethical and social complexities of new technologies…"* and the "we" of that position-holder is unclear.

23. The ESRC / ONS 2013 public dialogue findings, also concluded that participants did not have a clear or shared definition of 'socially beneficial'. Today, neither does government and it is hard to see how the Centre would define this, and not be bias towards the aims of the Centre, and towards the promotion of AI and its assumed entitlement to use data, over human rights which generally assumes a position of the right to privacy and autonomy.

24. For children this is enshrined in the UN Convention on the Rights of the Child, and we would expect the Centre to include this explicitly in its Terms of Reference or mandate.

25. The assumption that ethical approval from a committee always respects human dignity, autonomy, and human rights seen from the perspective of the panel, and that this makes a project morally acceptable to the people it is about, is a misnomer.

26. For example, the use of data linkage and machine learning across Health, Education, Social Care and Policing data to predict incidence of domestic violence and provide a heat map to police, is not how people expect their data, provided in confidence and trust, will be used. Women accessing the services of domestic violence charities, to whom we

---

[9] Guardian (Weale, S.) August 2018, *Student Loans Company 'spied on vulnerable students' social media'.* Students lost funding and dropped out of university despite no finding against them as part of an anti-fraud drive. https://www.theguardian.com/education/2018/aug/02/student-loans-company-spied-on-vulnerable-students-social-media

[10] defenddigitalme Timeline: school census expansion https://defenddigitalme.com/timeline-school-census/

spoke, pointed out that such use of resources in a risk-averse environment is highly likely to create higher levels of prediction than necessary. This could have adverse consequences for the application of police and support resources, and overburden rather than reduce workloads.

27. Emily Keddell, Senior Lecturer at the University of Otago in New Zealand, points out this is already very real in the area of child protection, saying of such systems, "*the data used to inform such models are incorrigibly suspect. Attempts to improve it lead to increasingly intrusive data use and challenges to legal equity. When such tools over-identify those least able to refute their 'high risk' label, we should all be concerned*."[11]

28. In the US, the Illinois Department of Children and Family Services shut down its Chicago algorithmic child abuse prevention program in December 2017, after the data mining software failed to flag at-risk children who died, while swamping caseworkers with alerts that thousands of others were at imminent risk of death. "*Predictive analytics [wasn't] predicting any of the bad cases*," Illinois Department of Children and Family Services director Beverly Walker told the Chicago Tribune newspaper. "*We are not doing the predictive analytics because it didn't seem to be predicting much.*"

29. Across the UK, public authorities already employ AI at scale, such as the AI at Xantura[12], without public debate, and that must be had alongside more transparency of the tools. This transparency of existing uses of data, bias, error rates, and preventing harm from the software design, should be a priority over expanding the use of AI and new uses of data.

30. Would the Centre shut things down if set up explicitly to promote '*innovation*'? If "*the Centre will not, itself regulate the use of data and AI*"? (1.9) What are its statutory powers?

31. It is possible that any future position and decisions may conflict with the bodies that do regulate data, notably the Office of the Information Commissioner, and if the Centre is on a statutory footing[13], whose view will win? How this is to be determined is not set out in the consultation, and remains unclear, but this is fundamental to understanding why a statutory footing may be necessary.

32. Trust in the work of the Centre will only be secure if there are no surprises in the purposes and outcomes of work undertaken.

33. There must be no perceived conflict of interests between the leadership, decision makers, and human rights of individuals and communities. This public perception is unavoidable if the leadership[14] of the Centre is under those who profit, or who have profited in the past, from the commercial or non-profit exploitation of public administrative

---

[11] Risk prediction tools in child welfare contexts: the devil in the detail http://www.husita.org/risk-prediction-tools-in-child-welfare-contexts-the-devil-in-the-detail/ Husita (accessed Sept 4, 2018)

[12] Xantura https://www.xantura.com/points-of-view/childrens-safeguarding-profiling-system (Sept 4, 2018)

[13] CDEO Consultation 1.11 "To enshrine and strengthen the independent advisory status of the Centre, we will seek to place it on a statutory footing as soon as possible"

[14] Public Accounts Committee: Dr Foster Intelligence: A joint venture between the Information Centre and Dr Foster LLP (p39) https://publications.parliament.uk/pa/cm200607/cmselect/cmpubacc/368/368.pdf

data, or if their own interests[15] should be seen to clash with any of the ethical implications of using big data.

34. In a survey[16] we commissioned in February 2018, 81% of 1,004 parents of children age 5-18 in state education, replied that parental consent should be required before sharing information about a child's special educational needs or a disability, with third parties such as researchers and commercial companies. Those who exploit national pupil data today, seem to have no interest in meeting that public expectation, or changing their own views.

35. There is a strong risk that the Centre will embed existing thinking of researchers, companies, and data users, but not the views of people whose opinions have never been taken into account, and do not know how the State uses their personal information today.

36. A thorough programme of public information about how government processes and retains public administrative datasets and a supporting programme of work to provide mechanisms for people to see their own data, and how to make Subject Access Requests, would not only restore a degree of legitimacy and underwrite the social license that is missing today, but provide an opportunity for the public to correct inaccurate data.

37. What the public wants first and foremost is that their data are used not by others for secondary uses, but are securely used under a duty of confidentiality, for their own care, such as "*sharing an individual's health data across different hospitals...the perception was that more data sharing could usefully be done within the NHS. There was also a strong feeling that personal health data are confidential, private and sensitive, and should not be shared outside secure, authorised bodies such as the NHS, and especially not with private companies such as employers, insurance providers and drug manufacturers. Mental health data was sometimes regarded as particularly personal and sensitive."*[17]

38. The Centre will not question its own existence. But to focus resources on re-use of data, while the primary purpose of its collection during the course of their everyday interaction with public services, is made lower priority, could be seen as inherently unethical by design. While it is right to be forward thinking, it cannot be at all costs in the present.

39. *"Many organisations in the United Kingdom are not taking advantage of existing technology, let alone ready to take advantage of new technology such as artificial intelligence."*[18]

---

defenddigitalme, September, 2018

---

[15] OPSN, Chaired by Roger Taylor, RSA Fellow and ex-Director of Research at Dr Foster, began as a project within the 2020 Public Services Trust chaired by 2020 Commissioner, Tim Kelsey. https://www.thersa.org/action-and-research/rsa-projects/public-services-and-communities-folder/open-public-services-network/about (accessed Sept 4, 2018)

[16] The State of Data 2018 survey commissioned by defenddigitalme, carried out by Survation [p34] https://survation.com/wp-content/uploads/2018/03/Defend-Digital-Me-Final-Tables-1.pdf

[17] Wellcome (2013) Summary Report of Qualitative Research into Public Attitudes to Personal Data and Linking Personal Data

[18] Select Committee on Artificial Intelligence, AI in the UK: ready, willing and Able? (March 2018) para 303 p94 (p96 of 183) https://publications.parliament.uk/pa/ld201719/ldselect/ldai/100/100.pdf