

## **National Statistician's Data Ethics Advisory Committee Minute**

Wednesday, 1 April 2020  
Meeting held via video conference

### **Present**

#### **Members**

Dame Moira Gibb (Chair)  
Mr Stephen Balchin  
Mr Rob Bumpstead  
Ms Vanessa Cuthill  
Mr Colin Godbold  
Ms Isabel Nisbet  
Ms Marion Oswald

#### **UK Statistics Authority**

Dr Emily Mason-Apps  
Dr Simon Whitworth

#### **Office for National Statistics**

Mr Peter Fullerton (for item 1)  
Dr Geum Young Min (for item 4)  
Ms Fiona Dawe (for item 5)  
Mr James McCrae (for item 5)  
Mr Billy Gazard (for item 6)

#### **Other**

Ms Jemma Gardner, Department for Education (for item 2)  
Mr Irfan Mohamed, Department for Education (for item 2)  
Ms Lia Sussman, Department for Education (for item 2)  
Mr Lewis Anderson, University of Oxford (for item 3)

#### **Apologies**

Dr Brent Mittelstadt  
Dr Emma Uprichard

### **1. Minutes and matters arising from the previous meeting**

- 1.1 The Chair welcomed members to the twentieth meeting of the National Statistician's Data Ethics Advisory Committee (NSDEC).
- 1.2 Members approved the minutes of the previous meeting.
- 1.3 Dr Whitworth updated the Committee with progress on actions from previous meetings. All actions were complete or in progress.
- 1.4 Mr Peter Fullerton presented the ONS Data Science Campus's approach to considering the ethics of the fast-paced work being conducted to produce

aggregate statistics that will inform the Government's response to COVID-19 in the UK.

- 1.5 Mr Fullerton explained that the Data Science Campus's work will only use anonymised or aggregate data, and will continue to adhere to ONS's strict security protocols for handling data, ensuring compliance with data protection and other relevant legislation.
- 1.6 The Committee acknowledged the requirement for this work to be conducted at pace, and welcomed the reassurance that ONS will continue to maintain strict security protocols and consider the ethics of their uses of data in the production of aggregate statistics to inform the Government's response to COVID-19.
- 1.7 The Committee recommended that the rapid ethical review process is limited to work that is being conducted to inform the Government's response to COVID-19, and that ONS evaluate this process once this period is over. Should any of this work continue after the COVID-19 pandemic then ONS should take this work through the usual ethical governance processes.
- 1.8 The Committee recommended that ONS continue to seek independent ethical advice when rapid ethical reviews highlight increased ethical risk, and offered to review work via correspondence in such instances.

## **2. Investigating vulnerabilities linked to serious youth violence using linked DfE and MoJ data [NSDEC(20)06]**

- 2.1 Jemma Gardner and Irfan Mohamed from the Department for Education (DfE) presented a proposal to use a linked Ministry of Justice (MoJ) and DfE dataset to explore the relationship between educational/social care factors on serious violence. This linked dataset includes the National Pupil Database (DfE), the Police National Computer (MoJ) and Criminal Courts data (MoJ). This project is being conducted in partnership with MoJ, the Cabinet Office/ Prime Minister's Implementation Unit, the Home Office, and the Office of the Children's Commissioner.
- 2.2 The Committee concluded that the purpose of this research as presented was not sufficiently clear, and that this application required major revisions. The Committee asked the researchers to address the following points raised during discussion in a revised application which the Committee would consider:
  - i. Clearly define the purpose of undertaking this research, outlining clear research aims and potential benefits for informing policy;
  - ii. Clearly outline the outputs that would result from each stage of this work, and the level at which these would be produced;
  - iii. Ensure that the use of data adheres to the principles of data minimisation and proportionality;
  - iv. Provide more information on the safeguards and procedures that will ensure the confidentiality and protection of sensitive and individualised data included in this linked dataset;

- v. Acknowledge and consider all the limitations and potential bias that exist in the crime data;
- vi. Carefully consider uses of terms such as 'predictive' and 'drivers', and remain cautious of implying causation when communicating findings or drawing conclusions;
- vii. Explore work that has already been conducted in this field, in particular lower levels of accuracy of machine learning for predicting life outcomes; and,
- viii. Provide clear assurance that this data will not be used for operational purposes, or to identify individuals or schools.

2.3 The Committee agreed to review a revised application at a future meeting.

2.4 **Action: Ms Jemma Gardner to provide a revised application to the Committee for consideration.**

**3. Parental separation and educational inequality: Any evidence for growing disadvantage or 'diverging destinies' across three recent cohorts? [NSDEC(20)07]**

3.1 Lewis Anderson from the University of Oxford presented a proposal to use the ONS Longitudinal Study (LS; 1971-2011) to explore the impact of parental education and parental separation on children's educational outcomes. The study will explore three successive cohorts of LS members aged 2-7 (and in a two-parent household) at baseline from the 1971, 1981, and 1991 LS cohorts. Each of these cohorts will then be followed up 10 years later (to measure whether a parental separation has occurred) and 20 years later (to measure educational outcomes).

3.2 The Committee raised the following points during discussion for the researcher to consider:

- i. Due to the nature of the LS datasets, there are a number of limitations to this work that should be explicitly considered. Examples include details about the health of parental relationships in instances where parents stay together, as well as a number of factors known to influence educational outcomes that this dataset may not enable the researcher to explore or account for in analysis;
- ii. Outputs and findings from this research should be clearly communicated to ensure that sound and proportionate conclusions are drawn;
- iii. The potential policies that this work could inform should be articulated better in the application;
- iv. Given the increase of alternative tertiary education schemes, the researcher should consider broadening the measure of educational outcomes, as well as being mindful that many individuals will not have completed tertiary education by the age of 21; and,
- v. The researcher should consider using the Scottish Longitudinal Study and the Northern Ireland Longitudinal Study to make this work UK wide.

3.3 The Committee approved this application, subject to Mr Anderson amending the application in light of the points discussed above.

3.4 **Action: Mr Lewis Anderson to update the application by addressing the points discussed in section 3.2 and provide this to the secretariat.**

**4. ONS Webscraping Policy [NSDEC(20)08]**

4.1 Geum Young Min presented a revised ONS Web Scraping Policy. This policy is applicable to all ONS staff activities involving web scraping of non-personal/non-identifiable data. This policy had been approved by the ONS Data Governance Committee who recommended that this Policy be shared with this Committee for independent comments and advice.

4.2 Dr Geum Young Min confirmed that all ONS staff who wish to web scrape data will be required to complete an ethical self-assessment and that these would be shared with the Data Ethics team. The team will refer projects to NSDEC in instances where ethical risks are identified as high.

4.3 Although the Committee acknowledged that the policy includes definitions of personal data as set out in the Data Protection Act 2018 1998 and Statistics and Registration Service Act 2007, the Committee recommended that Dr Min consider broadening the scope of this definition to include data that could be considered personal data but which is excluded from current definitions (for example, location data for an unidentified individual). The Committee also recommended that Dr Min include examples to illustrate this point.

4.4 The Committee were supportive of the revised policy, and agreed that it is the responsibility of the ONS Data Governance Committee to consider the acceptable level of risk regarding the outlined approach to web scraping within ONS.

4.5 **Action: Dr Min to consider the Committee's advice and update the policy accordingly.**

**5. ONS Sustainable Development Goals and Facebook Collaboration [NSDEC(20)09]**

5.1 Fiona Dawe presented a paper outlining plans for the ONS Sustainable Development Goals (SDG) team to explore the potential utility of using Facebook data to support the UK's reporting of the United Nations SDGs by filling data gaps and allowing insights around the goals and targets.

5.2 Facebook has made a specific commitment to the UN SDGs as part of the Data for Good programme, and approached the ONS SDG team with a proposal to collaborate. The SDG team are proposing to organise and participate in a 'hackathon' style workshop, which will enable a fuller assessment of the potential utility of this data before deciding on the efficacy of establishing a formal partnership.

5.3 Ms Dawe confirmed that the ONS SDG team will only enter into a collaboration with Facebook if the hackathon identifies that the Facebook data provides a clear and significant opportunity to improve the UK's SDG

reporting. Ms Dawe confirmed that the use of this data will not incur any cost to ONS.

- 5.4 The Committee were supportive of the steps being taken to consider the value and utility of this data before considering a formal collaboration, and encouraged the ONS SDG team to carefully consider the risks and benefits both before and after each stage of this potential collaboration.
- 5.5 The Committee recommended that the ONS SDG team ensure that they adhered to NSDEC's ethical principles when collecting and using Facebook data.
- 5.6 The Committee asked the ONS SDG team to provide an update on this work at a future meeting.
- 5.7 **Action: Ms Dawe to provide an update to the Committee at a future meeting.**

## **6. Update on gender identity questions in the Crime Survey**

- 6.1 Billy Gazard provided an update on the inclusion of gender identity questions in the self-completion module of the Crime Survey for England and Wales. These questions were added to the survey in October 2019.
- 6.2 Mr Gazard reported that response rates for the first quarter had been good, and no problems had been raised or identified by respondents or interviewers. The research team are confident that the inclusion of these questions will enable robust analysis and outputs that will be widely used by policymakers and service providers. The first publication from this work is due in early 2021.
- 6.3 The Committee welcomed this update, and also reiterated the importance of ensuring that the survey can be undertaken in the Welsh language.

## **7. Any other business**

- 7.1 There was no other business.