This note provides evidence from the Royal Statistical Society (RSS) to the consultation into the role and remit of the new Centre for Data Ethics and Innovation (CDEI) being conducted by the Department for Digital, Culture, Media, Sport (DCMS).

The RSS is a learned society and professional body for statisticians and data analysts. We have around 9,000 members worldwide - the majority of our membership is UK-based.

The RSS has campaigned for a Council for Data Ethics since our 2016 Data Manifesto, so we very much welcome the new Centre for Data Ethics and Innovation, but believe consideration is needed on the following five areas:

- **Nature and use of data essential to innovation**. Data quality, the origins of data, metadata and the computationally advanced modelling of data (including data linkage);
- **Health data.** The scope of how health data (and other data) is used more widely needs to be included in the wider strategic remit of the CDEI.
- **Skills needs for innovation**. There needs to be an understanding what the uses of data are and how to engage innovators and developers working in this area.
- **Data used for research.** Data-sharing across departmental or technical boundaries needs to be explored in greater detail with an emphasis on the benefit to society.
- **Personal data sharing.** Public opinion on personal data usage needs exploring in more detail with a need to balance privacy and the ability to use data for commercial innovations that would benefit society as a whole.

The independence of the CDEI also needs defining. The Centre's capability to challenge the data practices of departments and other data holding organisations is reliant on this.

Below are our responses to the consultation questions 1-8. A more extensive and evidenced account will be published separately on our website.

Q1 Do you agree with the proposed role and objectives for the Centre?

The RSS is largely supportive of the Centre's roles and objectives laid out in Section 2. In particular, we are pleased to hear the Centre will look to work closely with regulators to identify governance challenges and build on the best practice from public bodies and businesses. However, we wish to raise specific points on the following:

- **Regulation**. Point 2.2 examines the role of regulation, which is currently arranged around existing administrative structures. Departmental and technical boundaries should be challenged, and the Centre should look to expand who it works with. A small group of technology companies now hold immense amounts of data, and this will continue to grow.

The CDEI should work with competition bodies to consider how to ensure these firms do not become data monopolies and to consider ways in which the data they hold could more effectively be used for public interest purposes.

- **Ethics** are about norms and expectations and this might also mean co-regulating the sector and setting out what is expected, required and prohibited.
- **Data** is often understood as entries in a large spreadsheet. Whilst we are pleased to see that the CDEI supports the growth of responsible innovation, we would like to see an understanding of innovation reflected in more detail. Innovation requires thinking abstractly about mechanisms that produce this data. The CDEI should promote skills in this area, such as the analysis and interpretation of data.
- **Communication** by the CDEI should aim to explain how data is used and the benefits to the public. Research has shown significant public interest but displayed limited scope within existing background knowledge. It would be good to see an emphasis on communication from the CDEI about how data is used in the public domain and examples of where it can have a positive public impact.
- **Regulation** is currently arranged around administrative structures. The CDEI should foster understanding of innovation as a concept, and in practice. The current departmental and technical boundaries should be challenged in line with data and AI development.

Q2 How best can the Centre work with other institutions to ensure safe and ethical innovation in the use of data and AI?

We are pleased to see the commitment shown to working alongside other newly established organisations and regulators. The promise to work closely with research centres such as the Alan Turing Institute, Oxford Internet Institute and the Ada Lovelace Institute is welcomed. Professional bodies should also be engaged. RSS believes strongly in co-regulation and co-operation across the community.

In the specific area of algorithms, we recommend that the CDEI bring together specific regulators for this sector to review the use of algorithms. Sharing good practices and emerging issues should be encouraged and would lead to a comprehensive cross-sector understanding.
The CDEI must look outside the DCMS for guidance on how to structure its governance. Researchers in heath data, such as *Connected Health Cities*, are developing new models for governance, alongside the exemplar models that currently exist in research settings.

Although the international debate is discussed, the resulting international landscape is not described. This must include EU leadership, and leading ethical innovations from UK social research and official statistics.

Q3 What activities should the Centre undertake?

The CDEI should organise a diversified, cross-sector system of co-regulation, and work with learned societies, such as the RSS.

- **Data Monopolies.** A small group of technology companies now hold immense amounts of data, and this will continue to grow. The CDEI should work with competition bodies to consider

how to ensure these firms do not become data monopolies and to consider ways in which the data they hold could more effectively be used for public interest purposes[i].

- **Public engagement** must be a primary activity. Research shows that the public discriminates. Thus, the CDEI must prioritse them; their consent and endorsement of ethical norms is essential.
- **Usability of data** is important. Metadata standards for some data trusts should be set up and the CDEI should guard against manipulation of data. External regulation of data quality audits could potentially work well.
- **A new code of practice** could also fit a co-regulation model. The CDEI should share good practice with others and highlight emerging issues or incidences where there are gaps. Explicit whistleblowing policies might be required.
- **Cross-sector working**. Sectors and professions should be encouraged to set up their own working groups. The CDEI should promote and facilitate professional networking and sharing good practice.

Q4 Do you agree with the proposed areas and themes for the Centre to focus on? Within these or additional areas, where can the Centre add the most value?

We are pleased to see the proposals around data access. The areas of focus are important, and there is overlap and interplay between them.

The CDEI must address the delineation between health and other data. Innovation does not respect the siloes within which data are administered. Regulation should focus on the uses of data and not just on the origins.

The RSS are a representative body and we have received specific concerns regarding researchers experiencing restricted access to data (point 3.6) We look forward to being consulted on what the proposals for data sharing frameworks will look like and would welcome a role in providing evidence to help shape this.

- The impact of new uses of data is surprisingly hard to pin down. The CDEI should set up systems to collaborate with others to identify the impact of new uses of data and look for ways to integrate this. Analysis into the extent to how this impacts upon innovation and public benefit should be carried out and then communicated.
- Established specialist groups use evidence and draw conclusions on the innovative uses of data and the CDEI should ensure it has a targeted approach to engage with these groups.
- As data ownership is supplanted by Trusts licensing, the CDEI should ensure all uses are facilitated. Local governance models such as privacy commissions should be explored further.
- Synthetic data is emerging as an alternative in a variety of fields as a filter for information that would otherwise compromise the confidentiality of particular aspects of the data. In situations where researchers need greater access, the CDEI should support further professional accreditations and consider ways to enable data sharing.

---

[i] https://www.nature.com/articles/d41586-018-03912-z

- Time also affects the value and the sensitivity of data. Linking historical data to new data can bring new insights. History in data is difficult to destroy and more research in this area is needed by the CDEI.

Q5 What priority projects should the Centre aim to deliver in its first two years, according to the criteria set out above?

The project criteria do not currently relate to distinct projects. Here are some suggested activities that would meet with the criteria:

- The CDEI should commission a review of the **international** landscape and UK relation to it. This should be produced in consultation with other bodies such as the AI Council.
- The National Data Guardian delineated four data **uses**: service delivery; research; service improvement; and commercialisation. The CDEI should adopt or recommend alternatives to ensure clear separation.
- In reviewing the landscape, intuitive concepts emerge, such as data **ownership**. The CDEI might outline where these are flawed and make suitable recommendations to government.
- As data move to trusts, and algorithms become auditable and transparent, need for IP governance increases. The CDEI should review consequences and **patent** possibilities.
- When ethical **inadequacies** are identified a breakdown of trust can occur, curtailing legitimate uses. The CDEI should produce guidance on managing failures and engaging stakeholders to address public concerns and ensure lessons are learnt.
- Big data have profound implications on our society. **Interactions** between people and organisations can be mediated by algorithms and sensors. Specialist groups in the sector should be convened to share emerging issues. The CDEI might recommend attention in new areas.

Q6 What statutory powers does the Centre need?

Co-regulation across the many diverse sectors will need new powers that are appropriate to the size and stature of the organisation.
- The position of the CDEI in relation to other actors, including sector regulators, the Ada Lovelace Institute, and government bodies needs to be clearly defined and regularly reviewed.
- Similar bodies to the CDEI have already established external reviews by independent expert peers and the CDEI may find this format useful when concerns emerge. The CDEI must have the authority to refer them. The CDEI may need such review to take place where concerns emerge.
- The independence of the CDEI will be critical. The Lords Select Committee on Artificial Intelligence correctly advised that 'blanket regulation' should be avoided. There is the risk that giving the CDEI too many statutory powers mean it will cease to be a functionally independent body and becomes too intrinsically linked to Government. It must be able to object to public sector and business practices without fear or favour. Parliament should oversee this independence.

Q7 In what ways can the Centre most effectively engage stakeholders, experts and public?

The CDEI needs to be responsive and understood. It needs to engage and anticipate concerns.

- CDEI should draw on best practice where researchers have successfully engaged people about health and administrative data. Citizen juries have been effective - explaining data and use are essential. CDEI should commission such groups where appropriate.
- The CDEI should foster the formation of stakeholder groups. These should be sensitive to their sector's priorities and focus on the skills needs. The style and format of groups should be responsive to the diverse professional practices of their sector.
- The CDEI should be prepared to convene group members centrally. They can also share good practice and emerging issues as well as skills needs, in ethics and innovation.
- The CDEI should follow the public approach of the plain language summary required by METADAC. The lay public. should be the principal audience of every communication.
- Public discussion and media reporting of innovations, should be monitored by the CDEI and where misconceptions arise, the CDEI should be prepared to comment promptly.

Q8 How should the Centre deliver its recommendations to government?

Ethics is a matter of overriding public interest - recommendations to government must be made public immediately.

- There should be a parliamentary role for the CDEI's reporting mechanisms through sessional reports and regular hearings with the relevant select committee(s).
- Where a matter of national security is concerned, the CDEI should refer to the Intelligence and Security Committee of Parliament which has security clearance and sector familiarity.