

ADR England Public Insights Panel: Meeting 6 Summary Report

24 February 2025

The ADR England Public Insights Panel held its sixth meeting - and first session of 2025 - on 24 February 2025, 18:30-20:00, on MS Teams.

Contents

1. Attendance.....	1
2. Feedback from panel members	1
3. Forthcoming actions.....	3

1. Attendance

Facilitators: Tom Curran and Dan Clay (Basis Social), Shayda Kashef (Senior Public Engagement Manager, ADR UK), Sofi Nickson (Head of Research, Office for Statistics Regulation).

Public attendees: 10 out of 15 public attendees (referred to as panel members) were present.

2. Feedback from panel members

The purpose of the session was for panel members to give their views on priorities of the OSR's new strategy and share their thoughts on an infographic about synthetic data.

Sofi Nickson began by explaining what the OSR is and what they do, as well as what the OSR wants to know from the panel. Following this, the group was split into two breakout rooms to discuss three potential trade-offs, with Tom Curran and Dan Clay facilitating these discussions.

The panel's views were as follows:

Trade-off one: How statistics are made vs. How statistics are used

- The panel members in one of the breakout group prioritised 'how statistics are made', feeling that the importance of consistently getting the fundamentals right when creating statistics, outweighed the risk that some people may misinterpret or misuse them. For example, panelists suggested that, if statistics are not made correctly, even people trying to use statistics correctly may communicate incorrect information. By focusing on how statistics are made, OSR can at least reduce the risk that good faith actors have reliable statistics to rely on.

- The panel member in the other breakout group prioritised ‘how statistics are used’, feeling that the OSR has an important role to play in calling out misuse of statistics in the public eye. Furthermore, panelists have faith in the existing processes of filtering and checking that enable statistics to be made to a decently high standard, meaning the OSR should instead be able to focus their resources on improving the public’s understanding and interpretation of statistics.
- One panel member emphasised the need for the OSR to split their resources equally, if possible, across these two factors.

Trade-off two: Covering a small set of statistics vs. Covering as many as possible

- Panel members leant towards ‘covering a small set of statistics’ because they felt that it was most important to carry out rigorous quality checks on important sets of statistics, involve stakeholders in those checks, and pay great attention to detail where possible. For example, it would be better for the OSR to be able to guarantee that the statistics they have checked are correct, rather than to loosely check a lot of statistics. Finally, panelists felt that when looking closely at the details of statistics, new insights can be found, so it is important to look at statistics with depth rather than breadth.
- Many panel members sat in the middle, feeling strongly that there needs to be a balance between light touch checks and covering statistics that are genuinely important amongst wider society. For example, while it is important to cover small sets of statistics in detail, panelists were concerned that the OSR’s view of which statistics were “high profile” would not align with the opinions of the general public, making it important to cover a wide range of statistics that could be considered as such.
- One panel member leant towards ‘covering as many as possible’. They asserted that a small team could have a greater impact on removing statistics that are entirely wrong by carrying out light touch checks across as many statistics as possible, than to focus on a small set and leave lots of statistics completely unchecked.

Trade-off three: Regulating only official statistics produced by government vs. Encouraging people who produce other data to adopt OSR standards

- Panel members heavily leant towards ‘encouraging people who produce other data to adopt OSR standards.’ They noted that this would create a larger bank of good quality data and help push important data from outside of official statistics – for example, NHS dashboards, or statistics gathered using new technology – to a high standard. Finally, panel members believed that once other data had been brought to the OSR’s standards, more of the OSR’s resources and time spent on quality checks could be saved.
- One panel member prioritised ‘regulating and supporting only official statistics’, believing that because official government statistics are more widely known to the public, there are bigger implications in getting those statistics to a high standard, rather than unofficial, lesser-known statistics.

Following this, everyone had a short comfort break and then reconvened in the plenary session.

Tom Curran and Shayda Kashef then introduced the concept of Synthetic Data to the panel and invited them to share their thoughts on a related infographic.

Responses to synthetic data infographic:

- Many panel members were still unsure and confused as to what is meant by 'synthetic data', how this is different to 'anonymous data', and whether synthetic data is 'reliable'. Some panel members linked this to negative connotations around the word 'synthetic', such as 'poor quality' or 'not real', suggesting that an example might make things clearer.
- Some panel members felt the intended audience of the infographic was unclear, asking questions such as 'why is this relevant to me?'. One panellist initially assumed the infographic was intended for a researcher audience, as opposed to the general public. This may be due to the language used on the infographic, as several panelists felt that the infographic would be clear to researchers but would be bypassed or misunderstood by the public.
- Panel members associated the infographic with public health announcements during the COVID-19 pandemic; hospitals; privacy notices; research centres; and universities.
- Panel members suggested that the infographic could be improved by providing more information about what synthetic data is, and how and why it is used. Panelists would look to Google and other search engines to better understand synthetic data through a specific example. For example, one panellist felt they understood synthetic data better once they had used Google's AI summary to find out that synthetic data is used 'to validate mathematical models and to train machine learning models.'

Overall, the panel responded to the infographic with uncertainty regarding the definition, usage, and reliability of synthetic data, desiring more information.

Tom Curran wrapped up the session by explaining the relevance of Cardiff University as another stakeholder in this research and thanking the panel members for their time.

3. Forthcoming actions

- Basis Social will send out a homework task in the coming weeks for panel members to complete.
- The next session will be held in April 2025. Basis Social will canvass availability and preferences for the next meeting.