UNIVERSITY OF LIVERPOOL

Newcastle University

Centre for Data

ESRC Smart Data Research UK Strategic Advice Team

REPORT 4: SURVEY RESULTS

Rachel Franklin, Simeon Yates,
Elena Musi, Omar Guerrero, Seth Spielman,
Jeanette D'Arcy, Jessica Crosby

June 2025

DMSI
Digital Media and Society Institute

Centre for Urban & Regional Development Studies
CURDS

CONTENTS

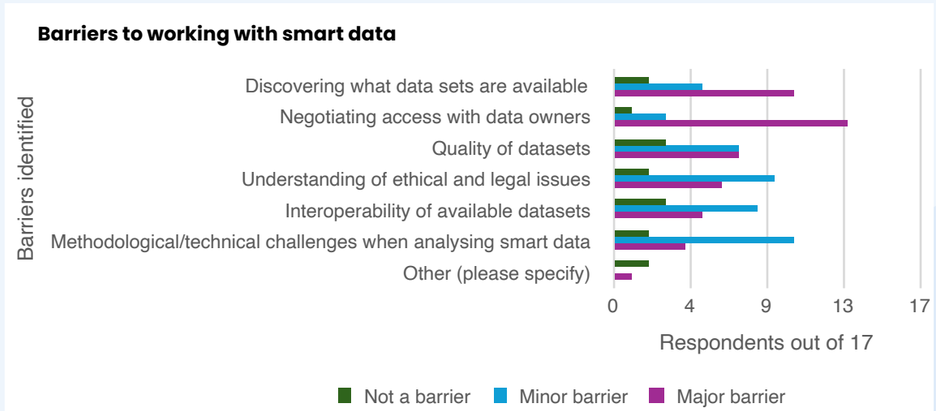
1. Executive Summary 4
2. Introduction 5
3. Methods 6
4. Findings 7
5. Delphi Responses 11
6. Limitations 14
7. Conclusions 15
8. Recommendations 16
7. References 17

Appendix One 18

Appendix Two 19

EXECUTIVE SUMMARY

- 22% of respondents **do not** currently work with smart data
  - In response to a question which offered statements to respondents and asked them to tick all that were relevant to them, 41% of those not currently working with smart data said they are unsure whether they would like to work with smart data and would like more information; 35% indicated that they would like to work with smart data but don't have the necessary skills; 35% indicated that they are very unlikely to work with smart data in the future; 12% said they intend to work with smart data in the next year.
  - For those not currently working with smart data, 'negotiating access with data owners' was the most commonly selected 'major barrier', followed by 'discovering what data sets are available':



- Both those who do and those who do not currently work with smart data identified 'negotiating access with data owners' as a **major barrier** to accessing or working with smart data
- Those working with smart data identified 'Rapid development of technologies' as the most common first choice when asked about **key barriers** to accessing and using smart data in terms of **risk and security**
- Most of those who indicated they do currently work with smart data have been working with it for **up to five years**, and mostly in the area of **research and/or teaching in academia**.
- 86% of respondents **somewhat or strongly agreed** with the statement 'Interview participants recommended that the ESRC create a resource centre, i.e., as 'first stop' that would provide guidance on, e.g., ethics, legalities and how to approach the use of smart data'.
- In terms of skills gaps, 'responsible research' was identified as the **most important skill area** for researchers, PhD students, and industry/government data providers. It was identified as second most important for those running data services, after 'data engineering'.
- 52 % of respondents agreed that **the likely future direction** for smart data in the UK is 'moving towards more open models of access', while 20% felt it will be 'moving towards more commercial models of access'.
- The survey had a completion rate of 43% (61 out of total 140)
- Respondents were **mostly male** (55%), mostly **41-50 years** of age (31%), from a **white British** background (82%)

The survey was carried out between March and August 2024.

INTRODUCTION

Research with smart data has the potential to improve lives, contribute to economic development, and provide valuable insights into the world we inhabit. Smart Data and related concepts such as the Internet of Things and Big Data are areas with large and increasing bodies of work in UK research. The use of smart data, its potential to contribute to society and the development of analytical methods in this area is often also concerned with the ethics and legalities of how smart data is accessed and used, with data justice and digital inequalities and with how policy and legislation can aid in efforts to utilise smart data effectively to tackle social, health and environmental issues. Smart data is used in work around smart cities, homes, agriculture and transport (see, e.g., Ghahremanlou et al., 2019; Laksch et al, 2021; Mahajan et al., 2021; Sovacool et al., 2020), as well as energy (e.g., Hargreaves et al, 2022) and healthcare (e.g., Brewer et al., 2023), among many other current and potential applications.

In terms of policy and legislation, data has become a key focus of government attention. The Online Safety Act recently came into effect, meaning greater regulation of online platforms, and the Data Protection and Digital Information Bill was moving through parliament, but at the time of writing the change of UK government from Conservative to Labour leaves this bill in question. The UK hosted the AI Safety summit 2023, with the aim of coordinating international efforts to consider and mitigate potential risks in development of AI technologies. While these documents are not solely concerned with smart data, they will all have implications for its accessibility and utilisation. In 2024 the government published the Smart Data Roadmap (Department for Business and Trade), outlining actions to be taken in key sectors of the UK.

The SDR UK investment comes at a time of increasing discussion around the power of data, and data infrastructure for innovative interventions in social and economic contexts. The Strategic Advice Team (SAT) aims to support SDR UK through a programme of investigation, to inventory and map the smart data research landscape via Delphi interviews and a questionnaire. This report focuses on the questionnaire portion of the project. The purpose of this survey was twofold. On the one hand, this survey followed on from a series of interviews with experts and was designed to provide feedback on initial findings in order to achieve consensus. On the other hand, early questions were designed to establish a baseline of information about who is using smart data for research in the UK, and the context(s) of that use.

METHODS

Design

The survey was designed by two members of the SAT team, using the online survey design tool Qualtrics. It was developed via an iterative process involving regular meetings and discussions and tested by the SAT team and members of the SDR UK programme before being administered (see Appendix 2 for sample questions).

The survey consisted of 29 questions (not including those which established consent). The survey was structured in two parts: the first asked respondents about demographics and baseline information (e.g., whether, how often and in what context(s) they use smart data), while the second part presented respondents with questions about their use of smart data for research, then statements regarding initial findings from the expert interviews with questions asking them to indicate levels of agreement. There was a separate route through the survey for those who indicated they do not currently work with smart data. Once they had indicated they do not currently work with smart data, they were asked questions about what barriers they faced, the potential for them to work with smart data in the future and what might enable this. They were not taken forward into the second part of the survey, as these questions were relevant to those currently working with smart data.

Ethics

Ethics approval was gained from the University of Liverpool. The opening text of the survey included information about the project and how respondents' data would be used, and written consent was sought from all respondents. This was a compulsory question (i.e., respondents could not proceed with the survey if they did not give consent). Ethical approval was necessary as we were also hoping to ask participants for contact email addresses, in order to share the data with the ESRC who potentially would want to revisit some of the findings with participants at a later date, for the purposes of project evaluation.

Participants and distribution

Respondents were sought that had experience in working with smart data for research, but in order to improve the representativeness of the sample we also sought responses from those who are not currently working with smart data but would like to. We sought out respondents through purposive and convenience sampling. Potential respondents were sent emails with information about the project and a link to the survey via relevant mailing lists, and relevant partner organizations were contacted and asked to share the survey link through their networks. The link was also publicised on the SAT social media sites (*X* and *LinkedIn*). The SDR UK team also shared the link through their relevant contact lists and through their SDR UK social media sites (*X* and *LinkedIn*).

Participants completed the survey online and there were 61 full responses recorded, all fully anonymised prior to analysis. Due to the purposive nature of the sampling process, respondents were largely academics working with smart data in a university context, but we aimed also to recruit both non-users and those working with smart data in other contexts (e.g., industry, government policy) in order to obtain a more representative assessment of the current smart data landscape.

Analysis

Survey results were analysed by the research team using Qualtrics, SPSS and Microsoft Excel. Based on the quantitative data gathered, we created graphs and visualisations, for example to show major versus minor barriers to working with smart data. Some questions, which had long-form text components, were analysed thematically by two researchers from the SAT team through a process of discussion and selection based on relevance, acumen, and insightfulness (see Appendix 3 for an example).

FINDINGS

Responses

There were a total of 140 respondents, but not all respondents completed all questions, and there were some drop-offs between initial questions and the main bulk of the questionnaire. 106 participants indicated their consent to take part in the study, with 96 completing questions about demographics, and 92 indicating in what context they would usually use smart data. 61 respondents completely finished the survey.

17 respondents who indicated they do not currently work with smart data completed the survey.

68 respondents answered the question about how long they have been working with smart data, dropping to 65 in the next question about how often they accessed smart data in the last year. The later Delphi questions were completed by between 42 and 47 respondents.

Respondent demographics

Respondents' answers indicated that those filling out the survey were mostly 41–50 years of age (31%), with 25% aged 31–40 years, and 26% aged 51–60. Most respondents identified as male (55%), with 38% identifying as female, 1% non-binary and 6% preferring not to say. In terms of ethnicity, most respondents were of a white British background, followed by other white backgrounds, with 82% respondents falling into these categories. The third most prominent background was Asian/Asian British (7%), with mixed/multiple ethnic groups at 5% and the lowest percentage of respondents from Black/Black British/Caribbean/African at 2%.

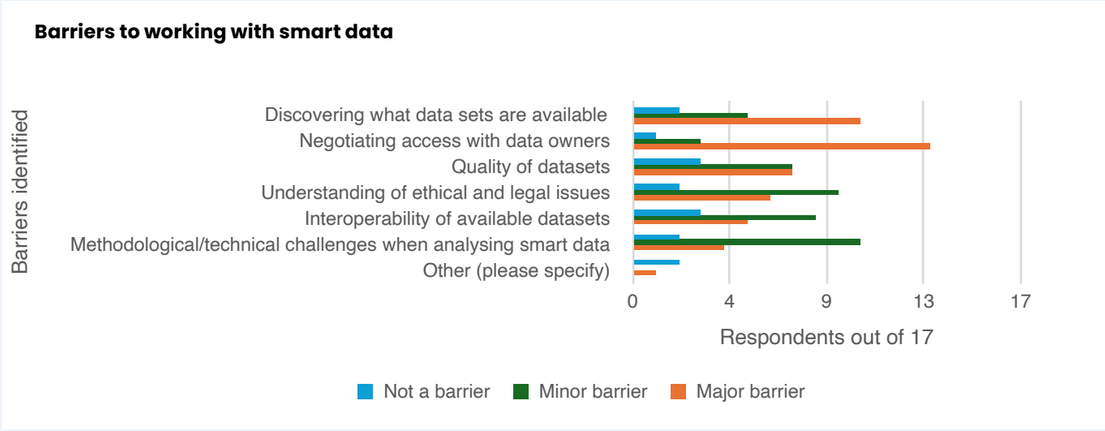
Experiences of accessing and using smart data

When asked in what context they would usually use smart data, most respondents used smart data in research and/or teaching in academia (60%). Some respondents used smart data in research for government policy (9%), or in industry (5%), with 5% choosing 'other'. Those who selected 'other' commented that they work with smart data in the context of communications, product development and design, and academic administration. 22% of respondents indicated that they do not currently work with smart data.

Those who indicated that they mainly used smart data in the context of research and/or teaching in academia were asked to indicate their disciplinary area and the type of data with which they mainly work. Areas with the most responses included Geography, Sociology/Social Science, Health and Psychology. In terms of the types of data being worked with, the largest categories were geospatial data, social media data, transport, financial and health (for the full list of disciplines and types of data, please see Appendices 1 and 2). Most respondents were working at either Early Career Researcher level (17 out of 37), or as an Established Researcher (17 out of 37). Three respondents identified themselves as PhD students (out of 37).

Those who indicated that they do not currently work with smart data were asked about the likelihood of their using smart data in the future, and what might be barriers to using it. Most respondents (41%) indicated that they were unsure whether they would like to work with smart data and would like more information. 'I am very unlikely to work with smart data in future' and 'I would like to work with smart data but

Figure 1: *Barriers to working with smart data for those not currently working with smart data.*



I don't have the necessary skills' equalled 35% responses each. Only 12% indicated that they intended to use smart data within the next year.

The 17 respondents who indicated they do not currently work with smart data were asked for more detail on what they felt were the main barriers to this, ranked from 'major barrier' to 'not a barrier'. 13 of 17 responses identified 'negotiating access with data owners' as a major barrier, followed by 'discovering what data sets are available' (10 out of 17 responses). Most respondents identified 'methodological/technical challenges when analysing smart data' as a minor barrier (10 out of 17 responses), with some identifying 'understanding of ethical and legal issues' as a minor barrier (9 out of 17 responses). Largest responses for 'not a barrier' were quality of data sets and interoperability of data sets (3 out of 17 responses each).

Respondents were asked to describe their level of confidence in working with smart data, rated on a 5-point Likert scale from 'extremely unconfident' to 'extremely confident', with the mid-point being 'neither confident nor unconfident'. The majority of respondents (51%) indicated that they were somewhat confident in using smart data, with 26% indicating they were either extremely or somewhat unconfident. 13% of respondents placed themselves at the upper value 'Extremely confident' while 9% placed themselves at the lower value of 'extremely unconfident'.

To examine the potential relationship between gender and levels of confidence among respondents, we cross-tabulated gender and level of confidence, shown in the table below:

Table 1: *Crosstabulation of gender and level of confidence.*

		How would you describe your level of confidence working with smart data?				
		Extremely unconfident	Somewhat unconfident	Neither confident nor unconfident	Somewhat confident	Extremely confident
What gender do you identify as?	Male	3.3%	10%	13.3%	53%	20%
	Female	13.3%	33.3%	6.7%	46.6%	0
	Non-binary	0	0	0	100%	0
	Prefer not to say	100%	0	0	0	0
TOTAL		8.5%	17%	10.6	51%	12.7

As shown in the table above, it appears that men identify as more confident in using smart data than women, with 22 of 30 respondents who identified as male saying they are somewhat or extremely confident, and no women identifying as extremely confident.

To expand on the barriers facing those who do not work with smart data, participants were asked to detail what might be limiting their confidence in utilising smart data, based on their own experiences in this area. Responses included:

- "Training and accessibility".
- "Lack of previous experience. Training tends to be descriptive, plus often focuses on 'what' you can do, rather than the 'why' necessary for policy".
- "I am at the early stages of developing my work so just need more time working with it to improve my confidence"
- "My programming skills are, ahem, *minimal*!"

In order to examine any relationship between age and levels of confidence among respondents, we performed a crosstabulation of age and level of confidence, shown in the table below:

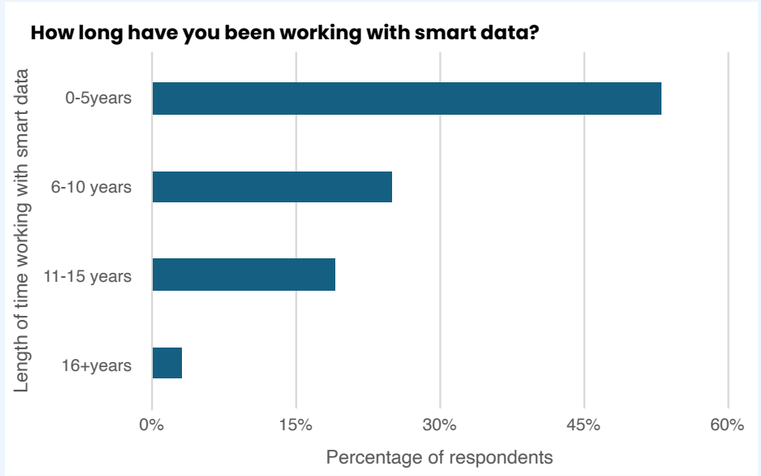
Table 2: *Crosstabulation of age and confidence level*

		How would you describe your level of confidence working with smart data?				
		Extremely unconfident	Somewhat unconfident	Neither confident nor unconfident	Somewhat confident	Extremely confident
Please indicate your age from the ranges below	20-30	0	33.3%	0	33.3%	33.3%
	31-40	23.7%	23.7%	0	46.2%	7.7%
	41-50	5.5%	16.7%	16.7%	50%	11.1%
	51-60	0	0	10%	70%	20%
	61+	0	33.3%	33.3%	33.3	0
TOTAL		8.5%	17%	10.6%	51%	12.7%

From the table above, it appears that confidence levels increase with age. 7 out of 13 (54%) 31–40 year-olds indicate they are either somewhat or extremely confident, compared to 11 out of 18 (61%) 41–50 year-olds and 9 out of 10 (90%) 51–60 year-olds. There were no respondents who identified as women that have been working with smart data for more than 16 years. As the above data (Table 1) showed that those identifying as women also indicated lower confidence than those identifying as men, it is possible that lower levels of confidence in women surveyed are correlated with their being at an earlier career stage.

Respondents who indicated that they work with smart data were asked how long they have been working with it, and how many times in the last year they accessed smart data. Most indicated they have been working with smart data between 0–5 years (36 out of 67), with some having 6–10 years' experience (17 out of 67) and fewer with 11–15 (13 out of 67) years of work in this area. Only 2 respondents had more than 16 years' experience of working with smart data. In terms of frequency of use, 22 % of respondents had accessed smart data more than 10 times in the last year. 22% had accessed smart data 4–6 times, while 49% of respondents had accessed smart data 0–3 times in the last year. Most of these respondents, when asked to describe their level of confidence working with smart data, chose 'somewhat confident' (51%), with 9% indicating they feel 'extremely unconfident'.

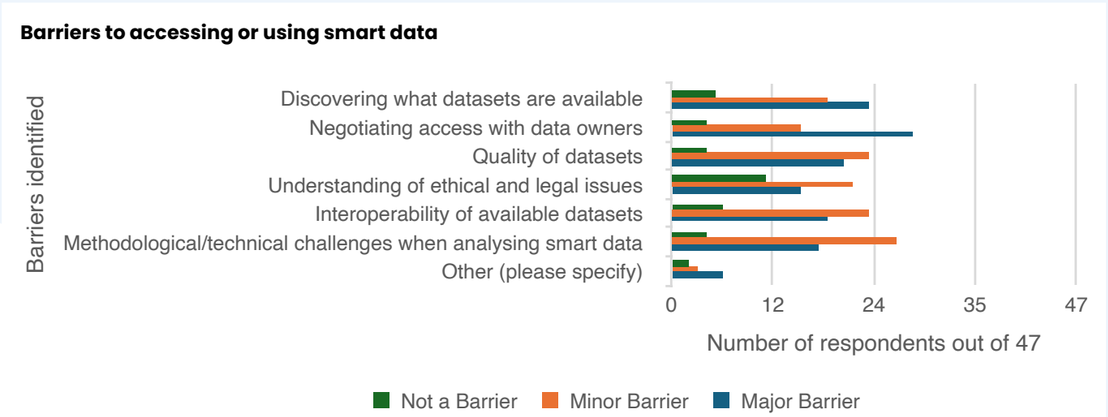
Figure 2: *How long have you been working with smart data?*



Respondents were also asked how often they have used smart data, alone or as part of a team, for research in the last year. Most respondents selected 1–3 times (52%), while 26% chose 4–6 times. Only 5% indicated they have used smart data for research 7–9 times in the last year, with 17% saying they have used smart data for research more than 10 times in the last year.

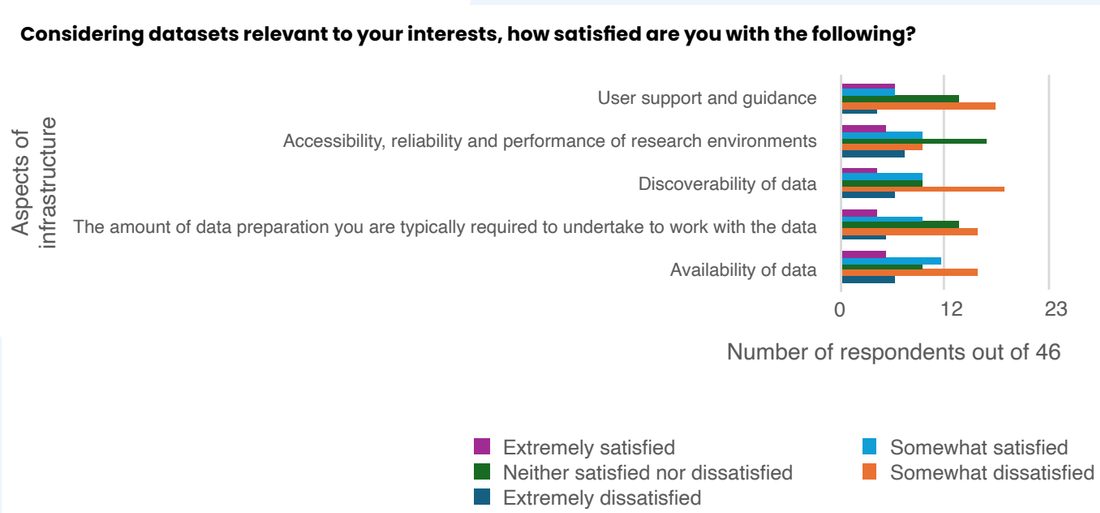
Respondents who said they work with smart data were also asked to think about any barriers they have faced when accessing or using it, and to rank options. 28 out of 45 identified 'negotiating access with data owners' as a major barrier, followed by 'discovering what data sets are available (23 out of 45). Around half of respondents identified three choices as a minor barrier: 'methodological/ethical challenges when analysing smart data' (24 out of 45), 'interoperability of datasets' (23 out of 45), and Quality of datasets (22 out of 45). The largest response for 'not a barrier' was 'understanding ethical and legal issues' (11 out of 45).

Figure 3: *Barriers to accessing or using smart data for those who currently work with smart data.*



Respondents who indicated they work with smart data were asked to consider datasets relevant to their interests, and rate how satisfied they were with elements of the smart data landscape. 24 out of 46 were extremely or somewhat dissatisfied with 'Discoverability of data'. 21 out of 46 indicated that both 'User support and guidance' and 'Availability of data' were extremely or somewhat dissatisfying.

Figure 4: *Levels of satisfaction in aspects of smart data infrastructure*



DELPHI RESPONSES

This section of the survey presented respondents with statements which summarised findings from the SAT team's interviews with experts and asked respondents to consider how far they agree. They were also given options suggested by participants in interviews and asked to rank them. This is intended as a step towards establishing a consensus viewpoint on key topics identified.

Statement: 'Interview participants recommended that the ESRC create a resource centre, i.e., as 'first stop' that would provide guidance on, e.g., ethics, legalities and how to approach the use of smart data'.

- 43% of respondents strongly agreed with this statement, and 43% somewhat agreed. Only 4% and 6% indicated that they somewhat or strongly disagree respectively.

Respondents were then asked to rank suggestions as to what should be provided by such a resource centre, from 1 – most important, to 6 – least important.

Table 3: *Ranking potential resources to be offered by 'first stop' resource centre*

	1 Most important	2	3	4	5	6 Least important
Standard guidelines for legal and/or ethical practice	28.5%	19%	21.4%	11.9%	16.7%	2.4%
A repository of up-to-date research in this area	11.9%	21.4%	7.1%	31%	11.9%	16.7%
Access and/or signposting to training and upskilling modules	21.4%	11.9%	28.6%	11.9%	21.4%	11.9%
Facilitation of a community forum	2.4%	9.5%	16.7%	11.9%	21.4%	38.1%
Access to a 'bank of experts' for advice on working with smart data	2.4%	9.5%	21.4%	23.8%	28.6%	21.4%
Data discovery portal	40.5%	28.6%	11.9%	9.5%	0%	9.5%

As shown in the above table, most respondents indicated that a data discovery portal will be the most important resource they want to see a resource centre deliver, with 29 out of 42 ranking this either 1 or 2. This was followed by standard guidelines for legal and/or ethical practice, which 12 out of 42 ranked first. The least important resource according to respondents was the facilitation of a community forum, which ranked at 5 or 6 for 25 out of 42 respondents.

Statement: *'Interview participants indicated the importance of building trusting, sustainable relationships between academia and industry in order to increase equitable access to smart data. One suggestion was the funding of 'liaison' posts for researchers as part of both industry teams and HE institutions, acting as a bridge between the two.'*

- 36% of respondents indicated that they thought this suggestion would be extremely effective, with 41% choosing 'somewhat effective'. 11% felt this would be somewhat ineffective, while only 5% felt it would be very ineffective.

Statement: *'A consensus which emerged from the interviews was that the general direction of smart data in the UK is towards more open access. However, some felt that it could also move towards more commercial models'.*

52% of respondents agreed that the likely future direction for smart data in the UK is 'moving towards more open models of access', while 20% felt it will be 'moving

towards more commercial models of access'. 15% chose 'don't know', while 13% chose 'other'. Those selecting 'other' were given the opportunity to explain, and gave the following replies:

- "Open access, through central funding"
- "Individual centric models"
- "A hybrid of models - commercial will kill off academic use unless funders support them (would they be deemed good value for money?)"
- "Bifurcation of access, with more open access to sample/pre-structured data, but only commercial models available for more bespoke research access"
- "A mix. Data should be open access wherever possible, but limiting to open access only means that some companies will never share their data with anyone"

In order to examine any relationship between how long respondents have been working with smart data, and whether they agree that future direction is towards open access, we performed a cross tabulation, represented in the following table:

Table 3: *Crosstabulation of time working with smart data and agreement with interview consensus*

		How long have you been working with smart data?			
		0-5 yrs	6-10 yrs	11-15 yrs	16+ yrs
Which of the following would you say is the most likely future direction for smart data in the UK?	Moving towards more open models of access	54.2%	16.6%	25%	4.2%
	Moving towards more commercial models of access	33.3%	55.5%	11.1%	0
	Don't know	85.7%	0	14.3%	0
	Other (please indicate)	0	50%	33.3%	16.7%
TOTAL		47.8%	26.8%	21.7%	4.3%

From this, it appears that those working with smart data for a shorter amount of time were also those who saw the future direction of smart data moving towards open access (13 out of 24 respondents).

Participants were also given several options identified by interviewees as key barriers to accessing and using smart data in terms of risk and security and asked to rank them from 1 - 'most problematic' to 7 - 'least problematic'. 'Rapid development of technologies' was the most common first choice (15 out of 46 scored it as 1), and 10 out of 46 scored this option at 3. 'Legal guidance' seems to have split opinion, as 10 out of 46 scored this option as 2, while 14 out of 44 scored it at 6.

Interviewees suggested equitable relationships between academia and industry might be further developed by funding 'liaison' posts, for researchers to embed themselves both in industry and Higher Education institutions, as a sort of 'bridging' role between the two. Participants were asked to scale how effective they believed liaison roles would be in strengthening relationships between academia and industry, with the scale running from 'very effective' to 'very ineffective'. 74% of participants agreed these roles would be very or somewhat effective. When asked to expand on their answers to this question, answers could again be scaled along effective or ineffective lines, such as:

Effective:

- "Having a well-identified "bridge" figure could be a cost-effective and efficient way to establish a maintain a trust relationship".
- "All projects, entities, etc. need someone to help bridge the different fields and perspectives. This is a very undervalued role, but a necessary one for anything interdisciplinary and across sectors to succeed".
- "A liaison post focused in a particular area (dataset / domain) can be effective in promoting collaboration, especially when linked to addressing key issues (data/ methodological/substantive)".

Ineffective:

- "Each organisation (potential data provider) will have somewhat particular data, ways of working, concerns, etc, so would be hard for such posts to achieve much generalisability".
- "Primarily my answer is given on previous experience of resourcing such posts. Generally they are recruited / funded at too low a level and should be experts at associate professor equivalent".
- "A post which is in two different teams/spaces rarely works well and one side normally dominates. Risk it could become an additional barrier to other researchers working closely with industry".

One respondent had an interesting suggestion regarding the remit of the liaison role, stating that rather than developing in-term posts, we could create "smart data champions", whose purview would be to engage a wide network of data providers to find solutions/mechanisms/enthusiasm etc. [that] could be more effective.

Skills gaps

In previous interviews with experts, skills gaps were a key theme emerging in the participant data. Survey respondents were therefore asked about what skills gaps they thought would need to be addressed by a smart data infrastructure programme, and to consider this question in relation to different roles in the smart data research landscape.

For PhD students, respondents felt that responsible research was the most important gap to be addressed (31 out of 44), followed by coding (for data science/ data analytics) (30) and qualitative methodologies (28).

For researchers, respondents also felt that responsible research was the most important gap to be addressed (37), followed by advanced statistical techniques (35), and coding (34). For industry/gov providers, respondents again identified responsible research as the key gap to address (26), followed by scicomms (22), and data engineering (21). For those running data services, respondents indicated that data engineering (31) was the most important gap, followed by responsible research (26), and AI/machine learning (19).

LIMITATIONS

As this work was focused on the experiences of smart data research for social good, purposive sampling produced a cohort of respondents that is largely within academia and so with limited representativeness of the broader smart data use landscape. Similarly, as a niche population, targeting those in the field of smart data research necessarily produces a relatively small sample size, affecting the ability to generalise results.

CONCLUSIONS

In terms of the demographics of respondents, the majority identified as men, were in the 41–50 age bracket and from a white British background. This could suggest that there is a lack of diversity in the field of smart data research, and/or that more research is needed to establish how diverse the smart data research landscape is. Respondents who identified as women indicated a lack of confidence working with smart data, with no women indicating they are 'extremely confident' and fewer indicating they are 'somewhat confident', compared to a significant proportion of men. However, the data also shows that confidence increases with the amount of time spent working with smart data, and no respondents who identified as women had been working with smart data for more than 16 years, whereas a small percentage of men did fall into this category.

In line with findings from other SAT work, a major barrier identified in the survey data is access and relationships with data owners. When asked to elaborate on what might help to improve their confidence working with smart data, respondents spoke about opportunities for training and developing skills, as well as more time to get 'hands on' experience through experimentation with datasets. The provision of a data discovery portal was considered important for any 'first stop' resource centre, as was standardised legal and ethical guidance. In line with interview findings, survey respondents seemed optimistic about the future of smart data being more open access. Survey respondents confirmed interview findings that suggested the creation of bridging roles between academia and industry, though not without significant caveats.

Survey findings suggest that smart data research is still an emergent field, with few indicating they have worked with smart data for more than 16 years, and not many using smart data on a regular basis for their research. This is reflected in responses to questions about infrastructure and skills gaps, where key areas of focus were data discoverability, along with user support and guidance for researchers. In line with interview findings, many respondents were concerned with the ethical implications of smart data research. Not only was this identified as an area of focus for any potential resource centre, but also as the main area of focus for any skills development and/or training.

RECOMMENDATIONS

In line with the above conclusions, the SAT recommends the following:

- The programme should encourage more diversity in the sector and boost the outputs of a diverse range of scholars.
- Further research should be encouraged on demographics among smart data researchers to establish a robust baseline of information on who is using smart data for research and in what ways.
- In terms of improving the confidence of those working with, or wanting to work with, smart data, structured training in data analysis should be prioritised, along with more informal opportunities to work with smart data in environments that encourage experimentation (e.g., the programme of Data Dives run by the SAT).
- Training should focus on responsible research and ethical practices, alongside the general data analysis skills identified in the recommendation above.
- The programme should continue to prioritise building relationships between industry, government, and private data owners and researchers.
- Further rounds of Delphi surveys could be undertaken to establish a robust consensus.

REFERENCES

Brewer, H.R., Hirst, Y., Chadeau-Hyam, M., Johnson, E., Sundar, S., and Flanagan, J.M. (2023)
Association Between Purchase of Over-the-Counter Medications and Ovarian Cancer Diagnosis in the Cancer Loyalty Card Study (CLOCS): Observational Case-Control Study
JMIR Public Health Surveill, 9. https://doi-org.10.2196/41762

Department for Business and Trade. (2024). *The Smart Data Roadmap*.

Hargreaves, N., Hargreaves, T., and Chilvers, J. (2022). Socially smart grids? A multi-criteria mapping of diverse stakeholder perspectives on smart energy futures in the United Kingdom. *Energy Research & Social Science, 90*. https://doi.org/10.1016/j.erss.2022.102610

Lida Ghahremanlou, Abdel-Rahman H Tawil, Paul Kearney, Hossein Nevisi, Xia Zhao, Ali Abdallah, A Survey of Open Data Platforms in Six UK Smart City Initiatives, *The Computer Journal*, Volume 62, Issue 7, July 2019, Pages 961–976, https://doi-org.liverpool.idm.oclc.org/10.1093/comjnl/bxy081

Laksch, J., Fernandes, E., & Borsato, M. (2021). Digitalization and Big data in smart farming – a review. *Journal of Management Analytics*, *8*(2), 333–349. https://doi-org.liverpool.idm.oclc.org/10.1080/23270012.2021.1897957

Mahajan, V., Kuehnel, N., Intzevidou, A., Cantelmo, G., Moeckel, R., & Antoniou, C. (2021). Data to the people: a review of public and proprietary data for transport models. *Transport Reviews*, *42*(4), 415–440. https://doi.org/10.1080/01441647.2021.1977414

Sovacool, B. K., and Furszyfer Del Rio, D. D. (2020). Smart home technologies in Europe: A critical review of concepts, benefits, risks and policies. *Renewable and Sustainable Energy Reviews, 120*. https://doi.org/10.1016/j.rser.2019.109663

APPENDIX ONE

Academic disciplinary areas of respondents

Science and Technology Studies (STS)
Transport
Economics
Urban planning
Computational social science
Geography
Human geography
Geocomputation
Psychology
Sociology/social sciences
Health
Law
Research software engineering
Financial data across a range of disciplines (RSE)
Fintech
Social media analytics
Arts and Humanities
Epidemiology
Information systems
Education
Communications
Remote sensing
Interdisciplinary

APPENDIX TWO

Types of data used by respondents

Social media
Transport
Administrative
Unemployment/employment
Geospatial
Environmental sensors
Invoicing data
Health (medical records, prescription/HES, administrative NHS, biosensing, eye tracking)
Financial
Housing (Airbnb)
Agricultural
Census (not smart data)
CCTV
Wearables
GPS traces
User logs
Retail footfall
Mobile phone (Apps)
Academic engagement (LMS)

DMSI
Digital Media and Society Institute

Department of Communication & Media
University of Liverpool
School of the Arts
19 Abercromby Square
Liverpool
L69 7ZG

Newcastle University
Centre for Data

Centre for Urban & Regional Development Studies
CURDS

THE ORIGINAL REDBRICK