

Big Data Network Phase 3: New and Emerging Forms of Data – Policy Demonstrator Projects

Call specification

Summary

Data can be thought of as the new raw material of the 21st century. Rigorous analysis of high-quality data can enable citizens to hold governments to account, drive improvements in public services by informing choice, and feed innovation and growth across sectors by providing a robust evidence base for policymaking and practice. The third phase of the ESRC's Big Data Network is focused on these New and Emerging Forms of Data.

Our ultimate ambition is to establish a new international, interdisciplinary 'Knowledge Hub' – drawing on existing and planned infrastructures and other investments made by the Research Councils and others to build a critical mass in capability, to access and analyse these new and emerging forms of data for policy-driven, impactful research in the social sciences. This hub will be founded on the principles of partnership-working, and will be focused on developing and delivering a high-quality service and data products which are of interest to and available for use by social research communities.

In order to inform the development of this larger investment, the ESRC is pleased to announce a call for **Policy Demonstrator Projects**. Projects funded under this call will demonstrate the potential value of New and Emerging Forms of Data to policymaking, either in their own right or when brought together with data from other sources. At the heart of these projects will be partnerships between academics and policymakers, and all proposals must include non-academic partner(s) from the public sector (eg government departments) who should play a significant role in the intellectual leadership and development of the project. They will also be required to make a cash or in-kind contribution. Proposals with international co-investigators are also welcome.

Applicants may request an ESRC contribution **up to a maximum of £200,000 (£250,000 full Economic Cost)** for a maximum duration of 12 months. All projects should start as soon as possible after the decision date, but no later than 14 February 2017. We may consider start dates up to 1 April 2017 by exception; please contact us if you anticipate this requirement. It is expected that approximately four projects will be funded.

These projects are being commissioned through a two-stage process.

At the first stage, Expressions of Interest are invited which provide details of the applicants, a summary of the proposed project, and an overview of the costs. **Expressions of**

Interest should be submitted via the online form available at <https://fs3.formsite.com/esrcforms/form74/index.html> by **16.00 on 14 October 2016**.

Applicants invited to submit full proposals will be informed by 21 October 2016, with **full proposals** to be submitted by **16.00 on 6 December 2016**. Funding decisions will be taken by a specially convened panel and will be expected no later than 27 January 2017.

Background and scope

Today's world is increasingly connected and digital; it is a world in which data are generated via a multitude of channels and devices on a second-by-second basis. In 2008, the number of 'things' connected to the internet (and each other) exceeded the number of people on the planet. By 2020 it's expected that there will be around 50 billion connected 'things', far outstripping the global population and all generating unprecedented volumes of data¹.

Since 2013 the ESRC has invested over £64 million capital funding in the Big Data Network, with phases seeking to facilitate access to a greater range and variety of data, stimulate innovative research using those data, and ensure that their full potential is explored and realised to the benefit of both researchers and data owners through development of new tools and methods for analysis; all the while maintaining the utmost consideration for ethics, privacy and research integrity.

- **Phase 1, the Administrative Data Research Network (ADRN)**, is working in partnership with government departments and national statistical authorities to facilitate access to administrative data that is routinely collected by government organisations, and helping accredited researchers to carry out social and economic research using linked, de-identified administrative data.
- **Phase 2, the Business and Local Government Data Research Centres**, were established to make data routinely collected by business and local government organisations accessible for academics in ways that prevent the identification of individuals, enabling them to undertake research that makes a difference: informing evidence-based policies and making businesses, voluntary bodies and other organisations more effective, as well as shaping wider society.
- **Phase 3 is in progress**. Having initially been focused on civil society data and social media data, the ESRC commissioned a round of **Civil Society Data Partnerships** with the aim of demonstrating the value of improved data infrastructure, enabling collection and analysis of data of interest to civil society organisations, and enabling the sector to better utilise its own data. In addition, three of the **Digital Economy Hubs (Horizon, SiDE, and dot.Rural)** were awarded enhancements, enabling them to build on their existing work around social media and conduct complementary activities.

Following extensive scoping, and in the context of a fast-moving data environment, it was agreed that the focus of **Phase 3** be **expanded from social media data to New and Emerging Forms of Data**, broadly categorised as follows:

¹ Cisco Internet Business Solutions Group, 2011 (<http://share.cisco.com/internet-of-things.html>)

- Internet data, derived from social media and other web-enabled interactions (including data gathered by connected people and devices, eg mobile devices, sensors, wearable technology, Internet of Things)
- Tracking data, monitoring the movement of people and objects (including GPS/geolocation data, traffic and other transport sensor data, CCTV etc)
- Satellite and aerial imagery (eg Google Earth, Landsat, infrared, radar mapping etc.).²

The ESRC also funds the UK Data Service (UKDS) which offers a wealth of experience in data-related archiving and access issues, and has more recently brought that experience into the new Big Data Network. It provides support and advice around pre-ingest activities (licencing issues for access to and use of data, rights and ethical issues), ingest activities (quality control, construction of metadata, migration to preservation formats), data management and archival storage, and data access including use of the UKDS's own secure data access facilities. In addition, we support methods and training activities including the National Centre for Research Methods, and we have recently commissioned a Centre for Doctoral Training focused on New and Emerging Forms of Data (see: www.esrc.ac.uk/skills-and-careers/studentships/postgraduate-strategy/centres-for-doctoral-training/).

The aim has been for all phases of the Big Data Network (including the current phase), their component parts, and other relevant infrastructures (data and otherwise) to work together to exploit synergies, share best practice, and overcome common challenges – to provide access to different forms of data, ensure that maximum impact is generated for the mutual benefit of data owners and researchers, and for the benefit of society and the economy more broadly.

Call details

The overarching aim of this call is to demonstrate/pilot the use of new and emerging forms of data for research in a policy context. All projects should seek to demonstrate the potential opportunities and challenges that these data (and, by extension, a large-scale infrastructure investment focused on them) might pose for policymakers, and indeed the broader policy-making landscape, for the purpose of enabling policy impact.

We are living in the age of the information economy, in a world where extraordinary amounts of data are constantly being generated. This increasing pervasiveness of technology and connectivity is enabling new ways of living and working, changing how social, economic, political, and cultural processes are created, and the resultant variety of new data are rapidly emerging as important to supplement, augment, and in some cases replace datasets collected by traditional means (eg surveys). Such data include (but are not limited to):

- Online search terms
- Website interactions (eg visit statistics, user-generated content, forums)
- Downloads (eg music, film, TV)
- Social media – text-based
- Social media – multimedia
- Blogs, forums and news sites (eg HuffPost, BuzzFeed, Reddit)

² From *New Data for Understanding the Human Condition*, OECD (2013) <http://www.oecd.org/sti/sci-tech/new-data-for-understanding-the-human-condition.htm>

- CCTV
- Sensors and connected devices (eg home energy sensors, traffic sensors, air quality monitors, wearable technologies)
- Mobile phone data
- GPS tracking data
- Satellite/aerial images
- Night-time visible radiation.

Many of these new and emerging forms of data arrive as live streams, and analysing them on a real-time or 'near real-time' basis can reveal phenomena as they unfold, prospectively informing decision-making, as well as potentially enabling more timely responses and immediate influence. Furthermore, data integration and the potential to augment traditional datasets by linking, calibrating and interoperating them with 'new' data offers opportunities to gain new insight and perspective on existing policy issues and to investigate potential issues in the context of an increasingly connected and digital world, including where these issues require joined-up approaches between multiple actors or stakeholders.

Projects funded under this call will explore how the new insights offered by these data can inform policymaking; in particular how they could make it easier to identify areas that require policy attention and better-targeted interventions.

The following broad policy areas have been identified as areas where new and emerging forms of data could have particular impact or relevance³:

- Health and wellbeing
- Work and skills
- Emergency/crisis response (including crime and policing)
- Law, justice and security.

All projects must employ at least one of the new and emerging forms of data from those listed above (alongside other datasets where appropriate), and applicants are encouraged to focus on one or more of the policy areas that have been identified. Proposals which focus on another policy area or areas will be accepted, provided that they are suitably evidenced and justified.

Under the terms of capital investment, project activities could include the demonstration or development of:

- New or existing infrastructures to enable identification of and access to and discoverability new and emerging forms of data 'at scale' – possibly exploring, for example, issues around data storage and re-use, data 'kitemarks', and researcher accreditation
- Tools and methods enabling more effective analysis of new and emerging forms of data, including in real- or near real-time
- Tools and methods enabling or testing the reproducibility of new and emerging forms of data
- Tools and methods enabling or testing the representativeness of new and emerging forms of data (including sampling methods)

³ ESRC 'New and Emerging Forms of Data for Policy' workshop, 20 July 2016

- Tools and methods which enable policymakers to test the potential for new and emerging forms of data from different sources or in different formats, to be linked, calibrated and interoperated (including with other datasets) in secure environments.

Other activities may be proposed, but are restricted to those which meet the requirements of capital funding. It is expected that successful proposals will see academic and non-academic partners working together to pilot the outputs of their projects in real-world policy contexts and seek to translate new insights into improved policy practices. All projects will also be expected to demonstrate the potential for scaling up.

Integral to all funded projects will be a commitment to the responsible use of data and ethical data practices. Applicants must include an exploration of any ethical implications that might arise from the proposed work, including issues around consent.

Funded projects should begin recording outputs within the first six months of the grant commencing. All projects will be required to submit a minimum of one case study demonstrating their work and the related policy implications by the end of the grant.

The start date for all proposals will be 14 February 2017. We may consider start dates up to 1 April 2017 by exception; please contact us if you anticipate this being the case for your proposal.

Eligibility

Applicants must be based at a research organisation eligible for Research Council funding; for a list of recognised organisations please see the RCUK website: www.rcuk.ac.uk/funding/eligibilityforrcs/.

All applicants will be expected to include letters of support at full proposal stage to demonstrate that there is genuine commitment from both participating ROs and non-academic partner organisations. Co-funding is not a mandatory requirement, however it is expected that all partners also make financial or in-kind contributions to the project, or a combination of the two. In-kind contributions might include, for example, allowing or facilitating access to data, facilitation of access to user communities, provision for non-academic researchers to analyse data, or use of facilities. There is no minimum or maximum level for partner contributions.

Letters of support from both co-investigator organisations and project partners must be signed on headed paper and dated no earlier than six months prior to the project start date. See the associated Je-S Guidance for further details.

Any proposals which do not meet these criteria will be rejected. For full guidance on the inclusion of non-academic co-investigators please see the ESRC website (<http://www.esrc.ac.uk/funding/guidance-for-applicants/inclusion-of-uk-business-third-sector-or-government-body-co-investigators-on-proposals/>) and associated Je-S Guidance.

In addition, we welcome the inclusion of international co-investigators and project partners. For information on the ESRC International Co-Investigators Policy, please see <http://www.esrc.ac.uk/funding/guidance-for-applicants/inclusion-of-international-co-investigators-on-proposals/>

Please note that it is the lead research organisation's responsibility to claim contributions from partners as specified in the proposal. Contributions from project partners, both financial and in-kind, should be detailed in the Je-S proposal under the 'project partners' heading. Any such contributions should not be included in the total requested from the ESRC.

If you are unsure about an organisation's eligibility under this call, please contact the ESRC in the first instance (see 'Contacts' below).

Eligible costs

Proposals should be costed under the normal full Economic Cost (fEC) rules. The funding limit for proposals is a maximum of £250,000 (100 per cent fEC). The ESRC will meet 80 per cent of these costs (up to £200,000), with the balance to be guaranteed by the host organisation. Funds may only be used for the activities listed under the 'Call details' section of this specification and not for substantive research. Associated studentships are not permissible under this call.

Please note that the capital nature of this funding requires that projects provide facilities, services and other support for data access and linkage, including purchase of datasets where appropriately justified, and with the aim of developing a resource which could have broader application and/or be of use to a wider community.

All proposals must include at least one non-academic co-investigator from the public sector, for example a government department or similar, who should contribute to the intellectual leadership of the project. Under the terms of the ESRC's policy, costs associated with non-academic co-investigators (regardless of how many are included) may not exceed 30 per cent of the overall cost of the grant. For further information on inclusion of non-academic co-investigators, please see <http://www.esrc.ac.uk/funding/guidance-for-applicants/inclusion-of-uk-business-third-sector-or-government-body-co-investigators-on-proposals/>.

How to apply

Stage I - Expressions of Interest

All potential applicants must submit an Expression of Interest providing the following details:

- Principal investigator
- Co-applicants/collaborators (including non-academic co-investigators)
- Types of data
- Policy areas
- Project summary
- Costs (approximate costs are acceptable at this stage)

All Expressions of Interest must be submitted via the online form available at <https://fs3.formsite.com/esrcforms/form74/index.html> by **16.00 on 14 October 2016**. Expressions of Interest sent by email will not be accepted.

Expressions of Interest will be considered by a subgroup, chaired by a member of the ESRC Capability Committee. This group will agree which applicants will be invited to submit full proposals based upon the following:

- Relevant background and expertise (of academic partners)
- Suitability of non-academic partners and wider collaboration
- Strategic fit – including policy relevance and data suitability.

Project costs will not be assessed at this stage.

Stage 2 – Full proposals

Shortlisted applicants will be invited to submit full proposals via Je-S (<https://je-s.rcuk.ac.uk/>) using the 'New and Emerging Forms of Data Policy Demonstrator Projects' call, and will be provided with appropriate guidance to do so. Full proposals must be submitted no later than **16.00 on 6 December 2016**.

All applicants (including non-academic co-investigators) must be registered with Je-S. Any co-investigators who are not yet Je-S registered may do so by going to the Je-S website (<https://je-s.rcuk.ac.uk/>) and selecting 'Create an Account'. Eligibility of non-academic co-investigators will not normally need to be checked if it is reasonably clear that they are appropriate to conduct the work specified in the proposal. If doubt is expressed during the commissioning process, checks will be carried out should the proposal receive a positive funding decision. Detailed information on the registration process will be provided in the associated Je-S Guidance.

Full proposals will be assessed by two members of a specially-convened commissioning panel against the following criteria:

- Innovation and originality
- Strategic fit (including policy relevance and data suitability)
- Value-added to existing UK social science research infrastructure
- Potential for future development of infrastructure and capacity from the demonstrator
- Potential to exploit synergies across the Big Data Network and wider UK social science data infrastructure
- Strength of partnership (eg co-production of knowledge, level of commitment/contribution made by partners)
- Expertise of key staff
- Overall value for money.

Please note that a single-stage review process will apply to full proposals, and therefore there will not be an opportunity for PI response.

Proposals will be ranked by the commissioning panel on the basis of excellence first and foremost; however, should this be equal and the level of funds available insufficient to fund all projects which fall above the quality threshold, the panel will be asked to consider the policy areas and types of data covered to ensure a diverse portfolio.

Commissioning timetable

- Call for Expressions of Interest opens – 7 September 2016
- Call for Expressions of Interest closes – 14 October 2016
- Full proposals invited – 21 October 2016
- Closing date for full proposals – 6 December 2016
- Panel meeting – 20 January 2017
- Applicants informed of decisions – w/c 23 January 2017
- Projects commence – 14 February 2017

Contacts

Any enquiries relating to the New and Emerging Forms of Data – Policy Demonstrator Projects call should be directed to nefd@esrc.ac.uk

Enquiries relating to technical aspects of the Je-S form should be directed to the Je-S Helpdesk (email: jeshelp@rcuk.ac.uk or telephone: 01793 444164)