Scoping the skills needs in the social sciences to support data-driven research skills across the academic career life course

Contents

1. Executive Summary

The Economic and Social Science Research Council (UKRI-ESRC) has commissioned this scoping study to map out the challenges, opportunities and areas of necessary change to existing UK data-driven research skills (DDRS) training provision in order to maximise the capabilities of future social scientists to contribute to and compete at the highest levels of international standards. The commissioned study also seeks to promote a model of life course DDRS training that enhances the UK's reputation as a global leader in the design, delivery, implementation and application of data-driven research skills that help address substantive questions of societal relevance and impact.

Data-driven research skills span qualitative and quantitative methodological applications and require a grounded knowledge of core principles in both domains as well as complementary mixed-methods approaches and applications. A DDRS training perspective that is future-proofed acknowledges that both quantitative and qualitative research methods accommodate data-driven research opportunities through which different data forms may be organised, analysed, interpreted and effectively communicated.

A life course model of DDRS training is proposed whereby it is acknowledged that DDRS training needs to move beyond a core focus on the postgraduate and very early postdoctoral career stages, to promotion and support of DDRS training across the academic life course, including during the mid- and later career stages.

The scoping study focused on addressing four primary question domains with the objective of informing recommendations specific to the design, implementation, support strategies, and future sustainability of a DDRS training framework that positions the UK at the international forefront of social science capability and capacity building in this area. These question domains are:
- What are social scientists' specific capability building and skills needs for data-driven research across the career life course?
- What interventions are required to enable social scientists to maintain DDRS across the career life course?
- What are the barriers to implementing and supporting a life course model of DDRS training?
- What are the possible mechanisms and strategies to effectively design, deliver and sustainably support a life course model of DDRS training?

The scoping study employed a complementary two-step methodology and approach to examining these questions and to facilitate primary implementation recommendations. First, an expert advisory group (the "Steering Group") was assembled, with expertise spanning multiple areas of social science and DDRS relevant knowledge and representing early to late career stage experience. Second, an independent research organisation (CFE Research) was commissioned to hold a series of sector-wide stakeholder workshops to understand how the vision for a life

course model of DDRS training might be achieved and how challenges and obstacles to attaining this vision might be overcome.

A suite of recommendations is provided that spans three proposed implementation themes (1) Implementation Models and infrastructure, (2) Developing a culture/environment for DDRS development across the career life course, and (3) Support for researchers.

Theme 1: Implementation Models and Infrastructure

- UKRI (including through the Research Excellence Framework, REF), the HEI sector and research funding bodies such as the Nuffield Foundation, Wellcome Trust and The Leverhulme Foundation should develop and adopt a model and infrastructure to design and deliver training/professional development in DDRS across the life course. This model and infrastructure should:
    - Enable the acquisition, consolidation and enhancement of qualitative, quantitative and mixed-methods data skills across all stages of career.
    - Be adaptive to new areas of data opportunity, challenge and innovation with training delivery responsive to new computational and digital skills developments, AI and machine learning, coding, programming, digital research methods, visualisation methods and other areas of data innovation.
    - Enhance knowledge and training in the areas of data linkage and related strategies to allow maximum utilisation of existing and future data infrastructure resources and opportunities.
    - Articulate the distinction and define the difference between DDRS and more generic research methods training skills needs.
    - Embed a plural model of methods training across all career stages.
    - Promote professional development (rather than 'training' alone)
    - Provide DDRS professional development and related training opportunities that will enable researchers to move across traditional discipline-based methods silos to allow a more multi-lingual model of DDRS training to operate across all stages of career.
    - Invest in and support the creation of an integrated resource platform for information concerning RO and commercial DDRS training opportunities.
    - Establish a DDRS training Centre of Excellence in collaboration with partners (e.g., Turing and/or Ada Lovelace Institute with additional engagement from business, industry, government and wider civil society).

Theme 2: Developing a Culture/Environment for Skills Development across the Life Course

- Funders, Research Organisations (ROs) and other influential stakeholders including national and local government, business and the third sector should collaborate to foster a research culture and environment that promotes the

development of DDRS across the life course by providing DDRS training opportunities, incentives, and recognition and reward for mid- and later-career researchers. Such an approach will enable:

- o The alignment of RO incentives to promote and reward life course DDRS training engagement (e.g. through recognition in cases for promotion and merit awards).
- o The establishment of research training funds and access opportunities for mid- and later-career researchers.
- o The specification of training and development expectations and opportunities through research grants.
- o Doctoral Training Partnerships (DTPs) to open up PGR specialist training to mid- and later-career researchers with appropriate funding/support aligned.
- o EDI responsive mechanisms to allow DDRS training engagement across mid- and later-career stages.

Theme 3: Support for Researchers

- To support researchers to upskill in DDR across the career life-course, funders and ROs should:
  - o Support the development of communities of practice that are cross-institutional and cross-sector and accessible across life course career stages.
  - o Provide funding mechanisms and incentives that encourage life course DDRS training engagement
  - o Specify training and development expectations and opportunities through research grants. For example, investments that develop novel DDRS methods and infrastructure could be encouraged to cascade training that is accessible and sustainable.
  - o Engage with ROs to jointly establish research training funds and access opportunities for mid- and later-career researchers.
  - o Link DDRS training opportunities and provision to research environment indicators (e.g., REF requirements) to encourage and reward life course DDRS training investment and future sustained impacts.

It is noted that implementing these recommendations will require further development in a number of areas. In particular, funders and other stakeholders should collaborate to ensure that the approach taken aligns with and supports the strategic ambitions presented in the Concordat to Support the Career Development of Researchers.

2. Introduction

Promoting strong data-driven research skills (DDRS) is a fundamental pillar of social science research training and capability building across all stages of academic career (doctoral, early career, mid-career and late career stages), and the bedrock of a sustainable social science research workforce for the future. ESRC recognises the importance of continually building and supporting DDRS among UK social scientists.

In the ESRC's 2019 Delivery Plan, the ambition to extend the investment ESRC makes to develop the skills and knowledge needed for social scientists to undertake high-quality, impactful data-driven research was outlined. Through this initial investment, ESRC endeavoured to ensure that social scientists have the data-driven research skills and capacity to fully utilise the increasing volume of large and complex data available for research purposes, and to maximise the value of existing investments in data and infrastructure. The importance of working across the breadth of quantitative and qualitative research methods and applications is recognised.

Building on this ambition, ESRC has commissioned the present study to identify existing strengths and gaps in DDRS training across all stages of a research career. (Note: this scoping study will specifically focus on career stages that extend beyond the period of postgraduate and doctoral training (see Tazzyman et al 2021)). The aim is to ensure that future policy and investment promotes and enhances DDRS among UK social scientists, and supports research leaders in maximising opportunities to effectively engage with data infrastructure in the UK and internationally across the academic career life course (see definition in Section 5).

This study builds on two recently published reports on DDRS and postgraduate training guidelines focusing on the doctoral career stage. (See Scoping the skills needs in the social sciences to support data-driven research – UKRI and ESRC postgraduate training and development guidelines – UKRI). However, ESRC recognises the importance of supporting DDRS training across the career life course, building and developing DDRS capability in the social sciences during and across career stages beyond the period of doctoral training. Additional research was therefore conducted to support this ambition.

The substantive focus of this scoping study is threefold. First, to highlight existing strengths and to identify gaps in the provision of methods training across early career (the period immediately following doctoral and post-doctoral training), mid-career and later career stages across all social science disciplines. Second, to recommend strategies that will allow current and future research leaders to enhance methodological skills that maximise opportunities to engage with the UK's world-leading and wider international data infrastructures. Third, to sustainably promote and enhance future data-driven research skills and capability building needs.

Data-driven Research Skills Scoping Study: The Need for a 'Life Course' Career Approach

There is a growing consensus on the necessity and importance of supporting and facilitating DDRS training across all stages of career. Presently, the greatest bulk of ESRC investment is concentrated during the doctoral and immediate post-doctoral stages of career. This study focuses primarily on career stages defined as early, mid- and later- career stages and addresses the following primary questions:

- What are social scientists' specific capability building and skills needs for data-driven research across the career life course?
- What interventions are required to enable social scientists to maintain DDRS across the career life course?
- What are the barriers to implementing and supporting a life course model of DDRS training?
- What are the possible mechanisms and strategies to effectively design, deliver and sustainably support a life course model of DDRS training?

Addressing these questions maps on to the UK Government's R&D People and Culture Strategy which sets out for the first time a whole R&D sector vision with an ambitious sets of actions in three primary areas; People, Culture and Talent, with an objective of attracting and retaining people of all ages and at all career stages in R&D roles.

3. Background and Context: Data-driven Research Skills for Social Science

Building on the implementation of ESRC's 2019 Delivery Plan, the ESRC has further specified strategic ambitions to build a world-leading DDRS training infrastructure in its 2022-25 Strategic Delivery Plan. It is noted that to achieve the primary ambitions outlined in this Delivery Plan, ESRC will continue to support the "health of all social science disciplines, maintaining a high-quality and diverse talent pipeline and a resilient, modern and innovative data infrastructure" (p. 4). Enhancing DDRS among UK social scientists across all stages of career is essential if this core objective is to be delivered. This section outlines some of the present and likely future drivers of DDRS and development needs and the potential benefits to UK social science and society.

a) Defining Data-driven Research Skills

The adjective data-driven means that research is determined by or dependent on the collection and/or analysis of data. It has acquired new significance given recent changes in the speed, scale and forms of data available for social science research. These changes – accelerated by digitalisation and increasingly ubiquitous computing – provide significant new opportunities for social science research. It has thus become important to identify the conditions that will ensure social scientists have the skills and capability to fully maximise the increasing volume of large and complex data spanning innovations in textual and visual methods, observational methods, digital technologies, machine learning and AI, as well as developments around large and complex data structures and big data.

b) Profiling the Wider Data Analytics Landscape

The nature, form and volume of data that are accessible to both quantitative and qualitative methodological applications have changed dramatically in the past decade. Digital data, textual and other forms of complex data have shifted the needs of DDRS training for UK social scientists. The widespread use of a multitude of digital and other data gathering resources and new technologies is fundamentally transforming the data landscape, with digital devices and applications becoming integral to everyday lives, and the amount of data being generated by individuals, groups and organisations expanding exponentially as a result. While complex in structure and prospect, these new and emerging forms of data have the potential to significantly enhance our knowledge and impact in multiple areas of social science application and societal relevance. The design requirements of future DDRS training in the UK therefore need to be responsive to the likely ongoing and rapidly changing data analytics landscape across multiple areas of social science relevance. It makes clear the importance of providing and supporting a life course model of future DDRS training to allow adaptation and capability to engage with these new (and continually evolving) forms of data.

c) Transcending the Qualitative/Quantitative Binary

As noted above, data driven research skills underpin the effective implementation of both qualitative and quantitative methodological approaches and require a grounded knowledge of core principles in both domains as well as complementary mixed-method approaches and applications. It is therefore essential that future DDRS training frameworks need to carefully embed a 'plural' methods training infrastructure that legislates against the historical quantitative versus qualitative single method 'siloing' and hence potential schism of skills training, recognising that methodological fluency across qualitative and quantitative domains will empower future data-driven research enquiry, capability and international competitiveness.

Further, the new data landscape associated with digitalisation and developments in computing provide an environment that poses new methodological challenges and opportunities. There is i) a need for training to recognise all stages of data collection and analysis, with ethical issues foregrounded. This may include participatory approaches as in citizen science; ii) a need for training to recognise the diversity of forms of data, including text and image as well as number; iii) a need for training to recognise the growing field of digital methods (e.g. Rogers 2019; Marres 2017) that cuts across quantitative and qualitative methodology and in addition requires specific digital skills – e.g. scraping URLs, extracting and analysing social media posts, images and videos; iv) there is a need for training that addresses specific computational skills – e.g. natural language processing.

d) ESRC: Maximising the Value of its Data and Infrastructure Investments

ESRC has made significant investment in DDRS and capacity building in the past. It is the largest public funder of social and economic research data in the UK and is recognised globally as a leader in social and economic data infrastructure investment and resource development.

In 2014, ESRC commissioned its Strategic Advisors for Research Resources to review the skills and capacity needed to develop within the social science research community to enable the social science community to maximise Big Data and other new forms of data and undertake research at the interface between the social and biological sciences. The review contributed to the evidence base informing two initiatives: i) steered studentships were allocated within Doctoral Training Partnerships (DTPs) and ii) the creation of two thematically focussed Centres for Doctoral Training (CDTs), one focussed on new and emerging forms of data (the Data Analytics and Society CDT) and another focussed on biosocial research (the Soc-B CDT) which draws on longitudinal data. DTP training guidelines and requirements have since been updated and enhanced in 2022 (Postgraduate Training and Development Guidelines (2022).

Additionally, ESRC has made strategic investments over the past two decades to build strong data skills capability. These include:

- Q-Step – an initiative co-funded with the Nuffield Foundation (and HEFCE until 2019, now Research England) to generate a step change in the teaching of quantitative skills in the social sciences, focussed at undergraduate level. Initially funded from 2013 to 2019 this initiative was extended to 2021. It provides a good example for understanding how skills training, when started early in the educational lifecycle, can lead to a cohort of graduates who can undertake data-driven research, and graduate into data-driven careers and advanced study. The findings of the [independent evaluation of the Q-Step programme](#) place emphasis on recommending Q-Step as a case study to universities teaching quantitative methods, and praise the value of relationships developed with industry (public, private and third sector) through work placements. This latter finding is reflected in the ESRC PhD review and in ESRC's strategic delivery plan (2022 – 2025).
- ESRC/Turing joint fellowship scheme – a core aim of the scheme was to build new interdisciplinary research capacity in data science and relevant social science disciplines. The Fellowships include support for post-doctoral researchers and PhD students.
- National Centre for Research Methods (NCRM) – ESRC established NCRM in 2004 to address long-recognised problems of methodological under-capacity in the UK social science research community. The Centre is currently funded until 2024 and has a focussed training remit and is responsible for the co-ordination of methods training funded by ESRC across its portfolio; acting as the first point of contact for social scientists seeking further information and training on research methods. Through its online portal, the Centre is intended to be a one-stop-shop, providing access to high-quality resources and training. It also aims to ensure a range of training provision and events are put in place, delivered both virtually and face-to-face, informed by an analysis of the training landscape. Whilst a general resource, NCRM offers training in some data skills competencies.
- ESRC was allocated funding from the National Productivity Investment Fund (NPIF) to support additional studentships in the areas of advanced quantitative methods (AQM) and data skills (38 in 2017/18 and 24 in 2018/19). Funding was also provided to support the Advanced Quantitative Methods Network (AQMeN) to map training provision in the social sciences in the areas of AQM and data skills. The team also piloted new training, designed with input from industry, in three areas where a gap in provision had been identified:
  - Data Wrangling
  - Predictive Data Analytics
  - Data Visualisation.

Continued investment in new and emerging technologies and forms of data has also been prioritised in the following:

- [AI review: Transforming our world with AI – UKRI](#)
- [National AI Strategy – GOV.UK (www.gov.uk)](#)
- [AI Roadmap – GOV.UK (www.gov.uk)](#)

- [AI Sector Deal – GOV.UK (www.gov.uk)](www.gov.uk)

Investment to build data skills capacity has also been made as part of ESRC's wider Data and Infrastructure Investments. For example, ESRC Business and Local Government Data Centres have developed and delivered a number of knowledge exchange, capacity building and training programmes such as:

- Master's Research Dissertation Programme: an annual national programme of industry Master's Dissertation projects (Consumer Data Research Centre)
- MSc Consumer Analytics and Marketing Strategy programme (Consumer Data Research Centre)
- ESRC Social Analytics Network (SASNet) – a programme of training and skills development activities (Business and Local Government Centre and the Urban Big Data Centre)
- Face-to-face training. Short introductory courses include: R; Stata; Social network analytics; QGIS; Geodemographic Segmentation; ArcGIS; Hadoop for Transport Informatics; training on connected vehicles; as well as a suite of advanced-level courses in analytics, visualisation and computational modelling.

Other relevant training initiatives are likely to have been established within individual research centres and large grants. However, this information is not captured systematically.

e) The Need for Ongoing DDRS Training, Capability and Capacity Building

Recognising the importance of continued investment in data and infrastructure, the ESRC recently set out a 5-year plan to capture, curate, combine and deliver data in a safe, secure and ethical way, while expanding the UK's existing national data resources and data facilities. The ESRC's 2022-27 Data Infrastructure Strategy is a framework which for the first time systematically captures how ESRC will invest in data infrastructure and associated capacity building, enabling social scientists to make the best use of resources, maximise the value of ESRC investments and build flexibility into the UK's data infrastructure portfolio. It is proposed that this infrastructure strategy will promote a joined-up landscape that is easily navigated by researchers and enables appropriate access to, and facilitates the use of, diverse data resources at a time of growing interest and demand from researchers, the public and others. This Strategy forms a key part of ESRC's broader 2022-25 Strategic Delivery Plan, which sets out ESRC's ambition for research that contributes to a more prosperous, healthy, sustainable and secure society. The Strategy also draws on priorities, approaches and information set out in the UKRI Strategy, National Data Strategy, and National AI Strategy. ESRC notes that investing in and maximising the value of its data and infrastructure has the power to be transformative, providing insight and evidence on a breadth and scale never previously deemed possible. DDRS training is integral to the delivery of this core ambition.

Several earlier reports commissioned by the ESRC highlight the need to address skills gaps in specific areas relevant to data-driven research skills:

- The Social Media for the Social Sciences study commissioned in 2016 (led by Professor Susan Halford) identified a skills gap in both the accessing of and use of social media data for social science research. The report recommended the development of training in the use of these resources across the talent pipeline.

- In 2017 ESRC asked for evidence on where knowledge and skills to improve the UK's social science capability and capacity should be targeted. This exercise identified skills needs in a number of areas relating to new and emerging forms of data including in the areas of: technologies; data; and computational science. Further information can be found in the ESRC summary report.

- The Longitudinal Studies Review conducted in 2017 recognised the value of high- quality data skills in the use of longitudinal studies and noted the importance of the availability of training in broader data skills. Subsequent work to map the skills needs to maximise the use of ESRC's longitudinal studies has highlighted a need for new training on data handling and data manipulation, including the creation of 'messy' datasets based on existing data that can be used for training.

In addition, a series of 'think pieces' were commissioned by ESRC in 2018 to inform the development of it's Delivery Plan priorities. Recommendations incorporated within these reports include:
- Making new forms of data and analytics become a basic part of social science training provision. They need to include not just computational skills but also data appraisal, data management and critical interpretation of analytical outcomes.
- Defining and delivering skills provision in emerging data, analytics, computational thinking and reproducible research
- An extension of CDT schemes supported by ESRC to provide more focused training in Data Science as well as greater coordination with existing Cross Council CDTs,
- The establishment of short term 'Mastering Data' Fellowships for larger cohorts of students to benefit from elements of either existing provision (CDT training modules) or new provision in core data science skills.

The Alan Turing Institute/ESRC Workshop: Social Data Science for Evidence Based Policy noted that there was an urgent need to define and deliver skills provision in emerging data, analytics, computational thinking and reproducible research. There is also a recognition that DDRS can help foster inter-disciplinary collaboration across and outside the social sciences. Initiatives such as those being undertaken by the Data Lab and the National Innovation Centre for Data provide one possible model, while also enabling the building of partnerships with – and transfer of skills to and from – the policy and industry sectors. The NCRM's 'Big qual analysis: Innovation in

method and pedagogy' project has also highlighted the need to develop pedagogy alongside cutting-edge methods, with a call for qualitative methods training and development in the use and application of large-scale qualitative data (Lewthwaite and Jamieson, 2019).

This is a growing, but still unbalanced, programme of skills provision across the academic sector that is focused on the early career stage rather than across the career life course and slightly skewed towards quantitative research rather than DDRS more broadly. Taken together with industry demands, this provides a compelling case for a more systematic approach to gathering evidence about what works at different education and career stages.

f)   Core Objectives of the Current DDRS Training Scoping Study

UKRI-ESRC commissioned this scoping study to identify existing strengths and gaps in DDRS skills training across the academic life course with the objective of sustainably meeting the future needs of UK social scientists. The core aim of the study is to ensure future policy and investment promote and enhance data-driven research skills among UK social scientists and support research leaders to maximise opportunities to engage with both the UK and wider international data infrastructure, while also equipping researchers across all stages of career with the skills to meet the needs of a rapidly evolving external data infrastructure environment and landscape (e.g. big data, AI and other areas of data-driven research skills development and opportunity). The core objectives of the scoping study therefore are to engage a systematic and methodologically guided approach:

- To review and recommend strategies to enhance the delivery of future data-driven research skills for UK social scientists
- To promote a model that moves beyond current siloed career stages of data-driven research skills training (e.g. early, mid-, later-career)
- To promote a life course model of data-driven research skills training that sustainably delivers future data-driven research skills needs and that is responsive to changing demands and environments.

4. Methodology and Approach

The methodology and approach implemented in addressing the core objectives of the scoping study comprise two principal stages. First, the DDRS Steering Group (see Appendix 1) conducted a systematic review of relevant literatures to define multiple areas of operational definition and relevance to the scoping study objectives (e.g. defining career stages, defining a life course DDRS training model, DDRS obstacles and opportunities, HEI priorities and other relevant areas of DDRS training review). Given that other reports addressed the needs of doctoral researchers, it was decided to focus on evidence linked to the provision of DDRS training at the early (post PhD training), mid-, and later career stages in order to better understand the needs of researchers at these career stages. Second, an independent research organisation, CFE Research, was commissioned to hold a series of stakeholder workshops to better understand how the vision for a life course model of DDRS training might be achieved, and how challenges and obstacles to attaining this vision might be overcome. The objectives of the workshops were set by the Steering Group which met approximately once monthly between September 2021 and December 2022.

Participants in the first two CFE led workshops (*n*=14) were invited to share their experiences of DDRS training in the current landscape as well as their ideas on how the vision for DDRS training could be achieved. The wide-ranging discussions explored (i) the skills needed by researchers in their mid-to-late careers, (ii) the strengths and weaknesses of current training provision, (iii) the challenges and barriers to training and development, (iv) strategies and interventions to build capacity for DDRS training amongst mid-to-late career researchers, and (v) mechanisms to ensure the successful implementation of a new DDRS life course framework.

Participants in a third workshop explored how the vision for a life course approach to DDRS training could be achieved, including how any challenges and barriers to the successful implementation of a new training and development framework could be overcome. Five senior academic stakeholders engaged in the discussion which covered (i) the extent to which there is likely to be support across the sector for the vision for DDRS training, (ii) the practical steps that need to be taken to deliver the vision, (iii) the main challenges and barriers to delivering the vision in the short term, (iv) strategies and solutions for overcoming the barriers, and (v) potential risks of implementing the DDRS training vision.

A synopsis of key discussion points and recommendations from the workshops is presented below.

5. Findings: Steering Group and CFE Workshops

i)   Steering Group Desk Based Research: Literature Review and Development of a Life course Framework

Steering Group members worked collectively and in work package subgroups between September 2021 and December 2022 to identify key themes relevant to implementing a life course DDRS training model. This involved a literature review and consultative meetings between members to review and integrate derived evidence and findings to help inform a life course DDRS training vision and proposed implementation framework. A synopsis of key areas of review and operational definition development is provided here.

Defining Career Stages

Historically, academic stage of career has been defined by early, mid- and later-career stages as if these stages are fixed or 'siloed' phases of career progression. DDRS training initiatives and opportunities have also been historically located at the doctoral career stage. This scoping study outlines the need for a life course model of DDRS training, thereby considering the skills needs of social scientists spanning early to late career stages. In this section we outline how the different career stages (beyond the period of doctoral training) have been initially defined for the purpose of this scoping study with the objective of highlighting some distinguishing features in relation to promoting future data-driven research skills development and continuous development of skills.

Early-Career Stage

The period immediately following doctoral training is an important phase of early career stage development. During this period, researchers' focus is on consolidating their PhD training and establishing an early research and expertise platform (e.g. developing a publication record, building a recognisable profile and engaging with relevant networks) to promote progression in their research area(s) and related academic/professional pathway trajectories (e.g., Lecturer to Senior Lecturer etc; Fellowship progression). Researchers at this stage of career are also typically seeking opportunities to lead and manage their own research programme. There is a large amount of literature focusing on DDRS training needs at this stage of career (e.g., Muniyappa, 2007).

Early career stage/status varies across funding and other relevant stakeholder communities (e.g., UKRI FLF) but represents a critical period during which DDRS may be promoted, consolidated and enhanced. Early training skills and acquired expertise (e.g. during the period of doctoral training) may also be harnessed during this period to begin to provide a return on training investment whereby this group can begin to facilitate and deliver training and skills development among new/incoming social science groups/cohorts. Recognising that this period can be up to 10 years post PhD offers the possibility for greater breadth of disciplinary representation (enhancing interdisciplinarity) and proficiency across qualitative and quantitative

methods domains to harness a 'workforce' that can both promote data-driven skills training and allow continued training and data-driven skills enhancement as a platform for transition and progression into mid and later career stages. It should be noted however that retention and skills building (e.g. supervisory experience) during this period are particular areas of concern where there is a preponderance and reliance on short-term research contract engagement.

Mid- and Later-Career Stages

There is limited consensus as to the entry and exit periods that define mid- and later-career stages. From the perspective of this scoping study, there is little in the literature specifically focusing on DDRS for those academics in mid- and later-career, but these career stages can be recognised as coming after an earlier, intensive training phase (doctoral and immediately post-doctoral). Nonetheless, mid- and later-career stages remain difficult to define and represent a long and ambiguous period. For example, Canale et al. (2013) highlight that by mid-career stage, academics and researchers may reach a plateau, where goals become unclear and support is lacking, meaning an overall trajectory can become vague. Similarly, Baldwin and Chang (2006) highlight that little is known about the mid-career stage in academia; they describe this stage as 'ill-defined', but refer to those positioned in mid-career stage as representing 'keystones':

> 'They fill essential instructional, programme development, administrative, and citizenship roles at their institutions. They form a bridge between faculty generations by mentoring new colleagues and assuming leadership duties as their senior colleagues move toward retirement. Mid-career faculty are key players as their institutions adapt in a time of continuous change' (Baldwin and Chang, 2006: 28).

Despite the lack of clarity, precise definitions, and the apparent ambiguity of these stages, for our purposes, mid- to later-career researchers will be defined as those who typically undertake self-directed research and can be categorised as independent researchers who have led competitively awarded research funding and related projects. Indicators that researchers are at this stage include:

- Experience of leading and managing teams

- Job packages combine teaching, research/scholarship, institutional and professional responsibilities within and beyond specific HEIs

- Evidence of continuity of engagement in research and teaching (as well as relevant research and related impacts) that position the researcher/educator on a continued academic or related professional pathway.

Further indicators may include:
- Extensive international research collaboration

- Extensive collaboration with external partners.

The identification of DDRS training needs for mid- to later-career researchers is often informal and driven by their programme of research. New skills are typically acquired through continuing professional development (CPD) and research-related activities and other informal routes as opposed to formal or dedicated periods of training/retraining. Mapping the skills needs and options for DDRS training during and throughout these periods has become increasingly important, including the mapping of continuity from early training into mid- and later-career training capability and a return to early career training for new and next generation researchers.

Disciplinary 'Silos'

In addition to noting stage of career silos in relation to accessing DDRS training options and opportunities, it is further noted that constituent social science disciplines often operate in theoretical and methodological disciplinary silos that preclude knowledge sharing in the area of data-driven research skills within and across stages of career. These disciplinary silos may be reinforced by processes such as the Research Excellence Framework (REF) and the apparent tension between disciplinary specialism and interdisciplinarity.

Mid-career stage researchers often have a methodological habitus which reflects disciplinary data skills and training. This can sometimes lead to siloed working with specific methods. Brannen's (2005) exploration of the possibilities in developing mixed-method approaches highlights that academics and researchers develop 'dispositions' and habits towards particular ways of doing things. As a result, researchers 'may lack the time and inclination to extend skills and interests in other directions and across the qualitative/quantitative divide' (Brannen, 2005: 174). Orientations and expertise in research and data skills are often rooted in the approaches a researcher was initially trained in. We can extend some of Brannen's (2005) observations to suggest that an individual's initial, early training may be an enabler or a barrier to retraining or developing new data-driven skills at later career stages. Researchers may be more or less likely to see the use or value in re-training or developing new and different data skills. Lukianova (2016) highlights that in adult training and education, the desire, willingness or need to train is influenced by motive and is informed by emotional factors, as well as the capacity to learn and the aims of learning. This connection between 'earlier' and 'later' suggests the need for a life course approach.

While progress has been made in this area at the doctoral and immediate post-doctoral training stages through DTP investment, further work and strategic investment is needed to enhance later career DDRS training and engagement opportunities, including in the area of interdisciplinary research.

University (HEI) Priorities and an Evolving Landscape

The role and significance of DDRS training beyond doctoral and immediate post-doctoral training stages within the context of University (HEI) priorities and a changing landscape has not previously been comprehensively reviewed. However, it

is widely recognised that career progression and reward pathways do not currently support mid- and later-career DDRS training engagement.

In the literature there is evidence and discussion of the constantly evolving academic and research landscape. This can create challenges for individual researchers, as well as for institutions more generally. It is within this context that DDRS training and skills development takes place. The Research Excellence Framework (REF) is a primary driver of University (HEI) business models but does not currently offer reward mechanisms for DDRS training engagement across all stages of career. Linking DDRS training opportunities and provision to research environment indicators may offer a mechanism to encourage, recognise and reward life course DDRS training opportunities.

Locke et al. (2018) highlight:
> 'perceived shifts are only just emerging in the UK and are due to a range of factors, including changes in policy and funding – and, in particular, the Research Excellence Framework – greater competition between HEIs, the increasing influence of global rankings, and developing international markets for academics'.

A comprehensive study of university 'research culture' commissioned by the Wellcome Trust (2020) made similar observations. They state that researchers report deep concerns about the sustainability of the current research culture in HEIs. This report suggests that too much emphasis on outputs is potentially undermining the quality of research, and that not enough is being done to address potential human costs of how research is produced.

Berg et al. (2016) summarise some of the realities facing many academic research staff, pointing to the measures of esteem academics and researchers must aim for to increase their 'future value', including:
> 'successful grant applications, peer reviewed publications (in journals with the 'right' impact factor), website blog posts, hits on their personalized socio-scholarly media websites (Academia.edu, ResearchGate, or Google Scholar), paper citations (via Thomson Reuters Citation Indexes, Elsevier Scopus database, or Google Scholar)' (Berg et al., 2016: 14).

Limitations of Later Career DDRS Training Opportunities

It is recognised that beyond the contexts of University (HEI) provided DDRS training infrastructure, primarily focused on doctoral and immediate post-doctoral training career stages, relatively limited formal training options exist for mid- and later-career stage researchers. While individual HEIs might offer specific training opportunities in areas aligned to core institutional strategic interests, this is highly variable across the sector with little externally funded incentives to encourage and support the development of DDRS training opportunities that meet the ongoing needs of mid- and later-career stage researchers. Addressing this limitation might involve an improved incentive model for HEI and other key stakeholders to engage (e.g., investigator support through grant funding mechanisms for mid- and later-career DDRS training).

Business and Other Stakeholder Access and Engagement Opportunities

A priority for the future of DDRS training in the UK is more effective engagement with business, industry, policy makers and other partners to maximise data-driven research skills training opportunities and the application of DDRS to addressing real-world, practical questions of relevance to UK social science and societal interests. Supporting more effective engagement with UK data infrastructure and specialist datasets that promote the interests and future engagement of business, industry and policy aligned partners across all stages of academic career will significantly enhance the utility of DDRS training and associated impacts.

Incentives to Engage in Later Career DDRS Training

A key challenge to implementing a life course model of DDRS training is the current limited access mechanisms, opportunities and incentives for researchers at later stages of career to engage in formal DDRS training (beyond individual interests and independent study).

This finding can be linked to the work of Crow (2020) who suggests research and academic career stages are forged in an 'evolving context (583) of demands'. Social science researchers must take chances and opportunities as they arise. He proposes serendipity plays an important role in the overall trajectory of careers, suggesting that attempts to consciously direct careers may be futile if they aren't contextualised or linked with existing opportunities or projects. Crow highlights that individual career trajectories within academia and research follow diverse paths respective to discipline, institution, funding environment, personality, and motivations. Crow also makes the point that many of the analogies we use to describe an academic or research career 'exaggerate the extent to which individuals control the direction taken by their careers' (2020: 588). Clearly multiple factors influence the overall direction of an individual's career, and the need or desire to learn new or develop existing data-driven research skills. Crow advises that 'it remains prudent advice to researchers not to be solely reliant on the winds of change to direct your career' (Woodthorpe cited by Crow, 2020: 590). We can infer from this that there may be no 'one size fits all' approach to training. To reflect the serendipitous nature of academic and researcher careers, training may need to be flexible and easily accessible rather than following a planned or fixed pattern. As such, funding bodies and other research institutions, including HEIs, must be able to respond adaptively and confidently to this 'opportunistic' landscape across career stages.

Like Crow, Whitchurch and Gordon (2009) highlight that the transforming landscape of HEIs means professional academic identities are becoming continuous and multiple, responding to and moving between different roles, tasks and workspaces. Millard (2016) similarly highlights that researchers in contemporary HE environments may not be tied to a specific institution, place or time, but experience mobility and develop skills for this. Both Whitchurch and Gordon, and Millard remind us that DDRS and training takes place within this context. However, there are many factors

that influence an individual's ability to learn. We must also ask whether individual researchers' and academics' motivations and values align with this necessity, as well as how researchers and academics can be best supported to make the most of opportunities. Reviewing grant funding mechanisms may be one immediate area of incentive promotion to help address this substantive limitation and obstacle to implementing and sustaining the proposed life course model of DDRS training (e.g. building in incentive-based funding options to engage in DDRS training opportunities).

It is important to acknowledge that there may be apprehension amongst mid- and later-career researchers in training in particular skillsets (see Leon and Brannen, 2015 for a humanities example). The 'fear' of approaching new fields or methodological approaches may impact an individual's willingness to participate in training. In addition, there is also a precedent in the doctoral research journey, where research increasingly points to the role of emotions in method learning (Nind et al., 2019). The literature on adult education and lifelong learning can help us understand this. As Gorodetskya (2012) suggests, continuous development and training throughout careers has a social-psychological effect because:

'The process of professional re-training and qualification improvement deals not only with mastering new knowledge and skills needed in the professional activity, but also personal socialization of a professional, i.e., some definite personal transformations, connected with interiorization of professional requirements, identification and adaptation' (Gorodetskya, 2012: 1).

Related to this, as highlighted above, developing new skillsets which are different from existing expertise will impact sense of purpose and academic or professional identity. Developing new skillsets is risky and there may be concerns of uncertainty and failure associated with this.

Supporting Access to Training and Developing Opportunities

Addressing the challenge of limited mid- and later- career DDRS training opportunities requires provision of future DDRS training access options and opportunities. This will also require an incentive model for HEIs and other key partners to allow engagement with these training options, and recognition of the ongoing need for DDRS training across the academic life course.

Barriers to accessing training can also include the complexity of language associated with certain data skillsets, as well as general awareness of training opportunities and resources required to complete them (this may include cost). Another problem related to access is adequacy. It may be that there is not enough specialised or advanced training available. Durrant et al (2015) asked ESRC Future Research Leaders about training provision, and the response implies that more training in advanced and specialised data skills is needed:

'When the FRLs were asked if they felt that the balance of training provision currently available to them between introductory, intermediate and advanced was about right, most felt it was not, with the majority wanting to see more advanced (specialised) training' (Durrant et al., 2015: 29; emphasis added).

This follows findings from Durrant and Lang (2004), which suggested that even in cases where academics and researchers identified training, it was very difficult to access these opportunities. Reasons given for this included: lack of funds, lack of time, lack of adequate opportunities/training and being unable to travel.

Training Needs

Although most institutions will have professional development reviews for mid- and later-career researchers, there is limited UK-based information on the training needs of this group of researchers. Currently it seems the most relevant and productive information available is from the NCRM training assessments and literature considering the research skills (not data-driven specific) training needs of early career researchers and those in more senior academic roles. This includes Durrant and Lang (2004), Wiles et al (2005) and Durrant et al (2015). Much of this literature focuses on particular methods related to social science research, in part reflecting the increasing availability of data from government and other surveys and data sources (with scant reference to non-quantitative forms of data). Some more recent work Kim, (2015) focuses on addressing 'the skills gaps in the rapidly evolving information professions' (in this case, around digital curation). Kim's research looked at research roles outside of the academy, with job advertisements scrutinised to uncover the skills needs of these roles. Kim's findings resulted in the development of four courses in digital curation which used project-based learning and active learning principles to develop new skills.

Sources of data for the analysis that resulted in course development to address skills needs included researchers' surveys, an NCRM events participants' survey, academic employers of researchers' survey and analysis of job ads for academic researcher posts. It appears the need for meaningful training needs analysis for mid- and later-career researchers links to skill shortages very directly. If DDRS training is to be embedded or even encouraged, it is important to understand where the shortages exist. For example, evidence suggests that, in the social sciences, there is the unique challenge relating to vastly different quantitative and qualitative training profiles in different disciplines. For example, Wiles et al (2009) highlight that in a UK study 78% of sociology PhDs, 77% of politics and 70% of social work students elect to use qualitative approaches in their PhD, whilst economics and psychology PhDs mainly use quantitative approaches. This may mean a large oversupply of qualitative researchers who do not possess quantitative skills, and likewise, a group of quantitative researchers with limited exposure to qualitative skills.

Areas of potential training need which are highlighted in the literature include;

- Using big data (including ethics of big data) (Couper, 2013; Ferretti et al., 2022)
- Data handling and management (Corti and Van den Eynden, 2015)
- Digital skills (Luhmann and Burghardt, 2022).

Durrant et al. (2015) and Durant and Lang (2004) conducted a review on training needs in advanced social science research methods. The 2004 review reported on a targeted questionnaire with 293 respondents from across the social sciences, consultation workshops and interviews with academics at Southampton University. Results highlight that there is a 'need for ongoing training throughout researchers' careers and that training needs identified were often discipline focussed'. More generally training needs were identified around managing junior staff, project management and learning programming languages, which may be extended to the need for a more plural model of cross-method language learning. One of the major findings from the 2015 paper is the call for a focus on training beyond the postdoctoral level (e.g., in the Recommendations 'training should be tailored toward mid and later-career researchers' (p. 40). Part of developing a future life course framework for DDRS training may require a consistent and thorough training needs analysis.

Changing Contexts and Future Sustainability

The need for cutting-edge DDRS training across all areas of social science application has never been greater. While quantitative, qualitative and mixed-methods methodological advances are occurring at a rapid rate, developments in the domains of AI, big data and other spheres of DDRS training relevance mean that the models of training provided to current and future UK social scientists have to be adaptive to future DDRS needs and responsive to future data-driven challenges and opportunities. This requires an approach that enables the acquisition of skills as part of a continuously developing professional identity. Training should be seen as an ongoing process – part of 'becoming researchers', while also recognising that there are different stages of development as well as different demands placed on researchers during those stages of development.

A useful model to conceptualise this process can be found in the work of Lave and Wenger (1991) and their notion of 'Communities of Practice' underpinned by the concepts 'legitimate peripheral participation' and 'situated learning'. For Lave and Wenger (1991) 'skill and knowledge emerge, not from the deliberate intentions of particular individuals, but through the co-engagement of the practice community as a whole' (Owen-Pugh 2007: 189). The aim of communities of practice is to 'build the organization's overall capacity to learn and innovate' (Wenger et al., 2002: 191). In Lave and Wenger's model, learners begin as legitimate peripheral participants with 'mastery' gradually achieved via shared goals and values and 'the transmission of community-specific knowledge and skills from one generation to another' (Unwin 2007: 110). The value of this model for understanding academic practice has been recognised before. For example, Mantai (2019) highlights the importance of participation in research communities for the continuing development of doctoral students and early career researchers. Degn et al. (2018) assessed four established research communities of practice. They noted the importance of 'socialisation' and that in communities of practice skills are developed through group dynamics. Degn et al. (2018) highlight that communities of practice act as a 'safe' space for 'novices' and early career researchers to learn and practice skills, and failure or mistakes are a necessary part of that process. It seems likely then that mid- and later-career

researchers would also benefit from the community of practice approach to skills acquisition. It has the potential to advance the recognition that academic careers require a continual process of training and capability enhancement, and that, for effective skill development to take place on an ongoing basis, established researchers and mid- and later career researchers require opportunities to become novices or legitimate peripheral participants in communities of practice to learn, or relearn, shared goals and values and achieve mastery of the new skills and techniques required for a changing data landscape.

It further suggests that training should be designed as part of a wider eco-system, which means that the interplay between University (HEI), government, business, industry, policy, funders and others will need to adapt if nationally relevant and internationally competitive DDRS are to be promoted and sustainably supported across all stages of career.

ii) Steering Group Contributions to Stakeholder Workshops

Specifying and Defining a Life Course DDRS Training Framework

Building on the findings of the desk-based stage of research just described, the Steering Group provided the following definition and models of life course training to CFE to inform the design of the workshops, comprising the second stage of the scoping study.

To support and enhance a sustainable DDRS training framework to meet future societal and social science needs, it is necessary to move away from a single stage model of skills training to a framework that recognises the importance of continuing skills building across stages of career. While a platform of fundamental DDRS training may be delivered during the doctoral and post-doctoral career stages, building and enhancing these skills through supported formal and informal DDRS training acquisition opportunities during early (period following doctoral and post-doctoral training), mid- and later-career stages is essential if the UK is to retain a reputation as a leader in DDRS training delivery and effective facilitation of engagement with the UK's world-leading and wider international data infrastructures is to be sustained.

There is a need for a shift in training needs to occur at two levels: (i) at the level of the individual researcher through opportunities to engage with and deliver training at all stages of career, and (ii) at the level of the primary delivery sites of DDRS training in the UK, that is, HEIs and other key disciplinary and professional partners/stakeholders. One proposal to be tested is that DDRS training be housed primarily in the academic/university sector and recognised as an 'ecosystem' where early training promotes mid-career skills building which facilitates later career skills enhancement and where these stages of data-driven skills building promote continued DDRS training and capability building for new and future cohorts of data-driven trained and skilled researchers. It is further to be explored how this ecosystem framework can sustain life course DDRS training progression while adapting to

external pressures and demands, allowing a sustainable framework both at the individual and system levels.

To begin to explore the value of this proposal, two models were reviewed: 1. Current DDRS training model and 2. A new DDRS life course training delivery model across stages of career, with a focus on ecosystem engagement.

DDRS Training – Current Model and Mechanisms of Delivery
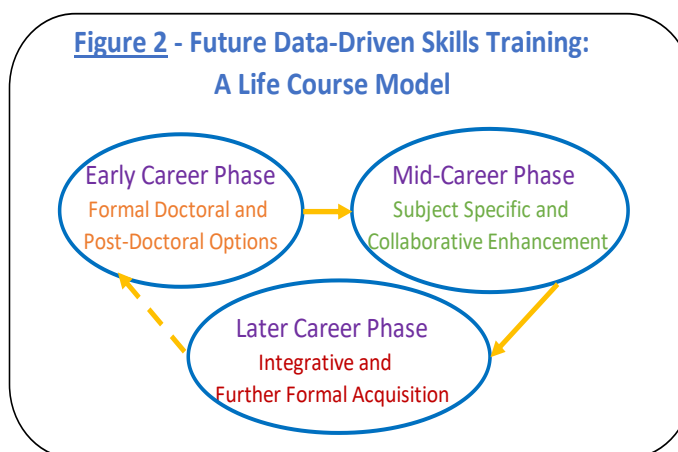
Current training provision across early, mid- and later-career stages may be grouped into three core domains (i) early career phase where DDRS training is delivered through formal doctoral and post-doctoral schemes. This period is marked by HEI and primary funder schemes where formal training is encouraged and facilitated through aligned funding schemes (e.g. post-doctoral training funding models), (ii) mid-career phase where formal training options are reduced in terms of breadth of focus and typically align more closely with subject specific and collaborative DDRS training needs, primarily delivered through informal mechanisms and *ad hoc* DDRS training opportunities (e.g. online training courses, subject specific conference training events; often at significant financial cost to the individual researcher), and (iii) later career phase where DDRS training is often regarded as unnecessary relative to stage of career needs with fewer options for training and limited financial support options to promote formal training opportunities. DDRS training at this stage of career may be regarded as adaptive to particular needs (e.g. grant funding focus relative to particular funding call specifics) and informally acquired through purpose-driven (e.g. research grant application) engagement or recruitment of earlier career researchers to facilitate purpose related DDRS needs. Across these stages of career, there is currently limited continuity of recognition for and formal mechanisms to help deliver DDRS training from early to mid-career and from mid- to later-career stages. There are few feedback loops or mechanisms of acquired life course DDRS training opportunities or mechanisms from those at a later stage of career to engage with those entering earlier career stages (see Figure 1) to help support life course continuity of DDRS training and adaptation to new knowledge and DDRS developments and innovations.

Figure 1 - Current Data-Driven Skills Training: Stage of Career 'Silos'

Early Career Phase — Formal Doctoral and Post-Doctoral Options

Mid-Career Phase — Subject Specific and Collaborative Engagement

Later Career Phase — Adaptive and Informal Acquisition
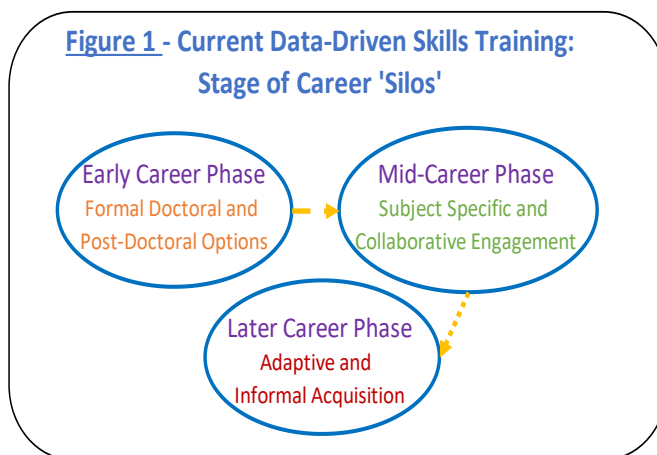
DDRS Training – Life Course Model and Mechanisms of Delivery

Building on the platform of DDRS training provided through existing (and newly developed) early career training options and opportunities (e.g. UKRI ESRC DTPs), it is possible that greater continuity for formal training options and opportunities could be supported such that a more cyclical model of DDRS training is created across the academic stage of career life course (including opportunities beyond academia). Fundamentally, in this proposed model DDRS training investment models accommodate greater continuity of training development and enhancement of skills as the researcher progresses from early career to later career stages. This will require a review of the obstacles and possible incentives to help promote continuity of DDRS training opportunities and engagement across life course stages of career (see previous section). Further, as depicted in Figure 2, delivery mechanisms that allow progression and continued engagement with DDRS training opportunities across the career-stage life course, promoting continued engagement with early career training infrastructure resources and new innovations will substantially enhance the DDRS training capability of UK social scientists in the future.

Figure 2 - Future Data-Driven Skills Training: A Life Course Model

- Early Career Phase — Formal Doctoral and Post-Doctoral Options
- Mid-Career Phase — Subject Specific and Collaborative Enhancement
- Later Career Phase — Integrative and Further Formal Acquisition

Building from the Steering Group desk-based research, review and recommendations, an independent research organisation, CFE Research, conducted two sets of stakeholder workshops to examine opportunities and challenges to developing and implementing the life course model of DDRS training framework that has been outlined.

CFE Workshops 1 and 2: A Stage of Career Perspective on DDRS Training

In this first set of workshops, participants represented early-, mid- and later career stages, with a broad representation of social science disciplines present. Participants were invited to share their experiences of DDRS training at their respective stages of career as well outline their ideas on how the proposed life course vision for DDRS training could be achieved. Discussions focused on: (i) the skills development needs by researchers in their mid-to-late careers, (ii) the strengths and limitations of current training provision, (iii) the challenges and barriers to training and development, (iv) strategies and interventions to build capacity for DDRS training amongst mid-to-late career researchers, and (v) mechanisms to ensure the successful implementation of a new DDRS life course framework.

a. Computational methods: Participants identified computational skills including coding as key skills for researchers to enable them to create, manipulate and analyse large or complex datasets. Participants indicated that they themselves, or the researchers they worked with, had upskilled in R and/or Python to enable them to do this. They also developed skills in the use of AI, machine learning and web scraping to facilitate DDR. Participants perceived that demand for these skills is likely to continue to increase, as is the application of computational methods to the manipulation and analysis of qualitative data.

b. Quantitative and qualitative methods: Participants noted that not only quantitative but also qualitative and mixed-methods engaged researchers require DDRS in order to create, manage and analyse large volumes of textual information and new forms of data. Proliferation in the formats and sources of qualitative data, including social media data, has led to the development of new methods for DDR including automated visual methods and digital social research.

c. Research design: Participants highlighted that by mid-to-late career, researchers have often developed specialist knowledge in their field and expertise in particular research methodologies. Participants recognised how important it is for researchers to also develop an appreciation of other methods and analytical approaches (quantitative and qualitative) throughout their career. Without this broader understanding, there is a risk that research designs will be driven by what the principal investigator (PI), in the case of research grant led investigation, knows and feels comfortable with, rather than the most appropriate or innovative approaches given the research question or 'grand challenge' being investigated. The participants highlighted the importance of methodological pluralism and the avoidance of siloed working in the context of DDRS training.

d. Research Leadership and management: Participants noted that as researchers progress in their career, they typically assume more responsibility for the strategic direction and management of DDR and undertake less of the practice. As such, the level of knowledge and skill required by researchers at different career stages varies. Those in the mid-to-late career stage may only require a high-level understanding of a particular method, programme or technique – sufficient to oversee a research project and supervise researchers (in an earlier career stage) who possess the requisite specialist knowledge and technical skills to undertake the research. In this context, those in their mid-to-later career need to develop the confidence and 'soft skills' to successfully bring together a multi-disciplinary team and facilitate collaborative working. However, a lack of detailed knowledge and skills in new and emerging techniques can present a challenge for mid-career academics when supervising students and post-doctoral researchers who have developed specialist skills during their doctoral training and/or post-doctoral stages of career. It is, therefore, important for

those in the mid-to-later career stages to have the opportunity to keep pace with developments in their field to enable them to continue to supervise students effectively.

e. Practical knowledge: Participants noted important issues relating to promoting future DDRS training opportunities. Specifically, participants highlighted that (i) researchers can now access a wide array of quantitative and qualitative secondary data resources, with requisite training in terms of data access and curation requirements essential, but not necessarily available, (ii) data-driven research can involve the use of commercially-sensitive and personal and special category data which is covered by the General Data Protection Regulations (GDPR), necessitating the importance of training in data protection and data security protocols, and (iii) the importance of research ethics training, with participants noting that while by mid-career, most researchers have a good understanding of the procedures that need to be followed to obtain ethical clearance for their research, there is sometimes a gap in understanding about how to apply ethical principles in practice (e.g., social media data). Participants observed that researchers need to understand the ethical considerations of methodologies that involve the use of personal or sensitive information and have the skills to assess the integrity and potential impact of their approach on data subjects to ensure they produce robust and ethically sound research.

2) Strengths and Weaknesses of Current Training Provision

f. Strengths: Participants identified a number of strengths in relation to current DDRS training infrastructure and options. Examples tended to focus on individuals and individual academic institutions, although participants noted that access to digital training options has improved as a result of Covid-19 as providers have shifted their delivery online. The discussion highlighted the range of training options available for those seeking to develop technical skills such as R or Python, particularly in the commercial sector. These providers were perceived to offer high-quality, cost-effective training that was suitable for people at different career stages. Participants were also aware of relevant DDRS training offered by the National Centre for Research Methods (NCRM). HEIs were also recognised as a source of expertise in DDRS training and pockets of good practice were identified. It was observed that the development of topic specific training by individual HEIs was leading to duplication across the sector as well as variability in the quality and content of provision between institutions.

g. Weaknesses and gaps: Participants questioned the extent to which training across both quantitative and qualitative domains is sufficiently tailored and pitched at an appropriate level for those in their mid-to-late career, although this was particularly highlighted in relation to qualitative approaches. There was a consensus that training is often geared towards the development of specialist technical and predominantly quantitative skills. It was also noted that current training

provision tends to concentrate on 'how' to do something rather than 'why'. It was noted that it is important for researchers, and those responsible for research design in the mid- and later-career stages, to be able to recognise under what circumstances it is practical and appropriate to use DDRS and to be able to provide a rationale for using specific methods (i.e., what is the question?). Participants also noted that the demand for interdisciplinary research is increasing and that this brings with it its own challenges in relation to DDRS training. While most participants acknowledged the importance of leadership and management skills for DDRS training, there was some disparity in their views on current training provision. While some perceived there to be a wealth of leadership and management training, others identified a lack of suitable training in these 'softer' skills relative to the volume of specialist and technical skills training available.

3) Challenges and Barriers to Developing DDRS Training

a. Time: participants agreed that lack of time to both source and undertake training and development is a primary barrier, particularly in the mid- and late career stages. In their experience, the onus is often placed on the individual to identify training. The lack of a framework or signposting to quality-assured provision means a significant amount of time is invested in searching for suitable provision. When a suitable option/course is identified, those in their mid-to-late career stage often find it challenging to carve out the time to complete the training alongside their other commitments. There were numerous examples of researchers undertaking training in their own time and at their own expense.

b. Funding: Lack of funding was frequently mentioned as a barrier to DDRS training. Those in more senior positions with financial responsibilities have the flexibility to fund training for themselves and for the staff in their department or centre. However, those in their mid-career who are not grant or budget holders are unable to easily access funding for training, either via institutional staff development budgets or via external sources. As a consequence, some researchers have funded their own development.

c. Incentives: Participants highlighted the importance of locating any proposed DDRS training recommendations within the current (and expected) HEI research assessment (Research Excellence Framework, REF) and external funding landscape (primarily grant funding). In terms of external drivers, it was noted that current metrics linked to REF are not favourable to promoting continued career or life course engagement with DDRS training. Current grant funding models were also noted as limiting, and as failing to promote or incentivise mid- and later-career DDRS training. HEI practice and culture was also highlighted as a limiting factor in terms of promoting a life course model of DDRS training. It was particularly noted that institutional practices, including criteria for progression and promotion, reflect the external

landscape and typically reward staff for traditional research outputs. The development of skills (including DDRS or leadership and management) is not typically recognised and, as such, is neither valued nor rewarded. Furthermore, the research environment in the context of academia is often characterised by 'soft money' and fixed-term contracts which negatively impact staff development and staff retention.

d. Support for the DDRS training vision: Participants reported that the current DDRS training model, with its focus on discipline specific training delivered primarily during doctoral training and the early career stage, is outmoded and no longer fit for purpose. There was broad support for the vision set out by the Steering Group amongst participants, who agreed that a new life course approach to DDRS training is needed to ensure researchers are equipped with the knowledge and skills to adapt and keep pace with the demands of an evolving research landscape throughout their careers. Participants recognised that there may be resistance from some within the sector to the concept of training and continuing professional development during the mid- and later career stages. To achieve buy-in to a new approach, it was highlighted that it will be important to couch the vision as a response to the changing nature of the world of research, rather than a deficit in the skills of those in their mid and late careers. A life course approach to learning and skills development has been adopted in the commercial sector and embedded in the culture of most organisations. Learning from peers and colleagues, irrespective of their career stage, and employing experts to deliver training and coaching to staff are commonplace. It was suggested that greater co-operation, including the sharing of practices, between the commercial and academic research sectors could help to overcome barriers and support the successful implementation of the vision within the HEI sector.

CFE Workshop 3: An Implementation Perspective

In a third workshop that involved senior HEI and other participants, discussion focused on how the vision for a life course approach to DDRS training could be achieved, including how any challenges and barriers to the successful implementation of a new training and development framework could be overcome. Five senior academic stakeholders engaged in the discussion which covered the extent to which there is likely to be support across the sector for the vision for DDRS training, the practical steps that need to be taken to deliver the vision, the main challenges and barriers to delivering the vision in the short term, strategies and solutions for overcoming the barriers, and potential risks of implementing the DDRS training vision. Discussion and recommendations focused on the following core topics.

1) Overcoming Barriers to DDRS Development

Participants identified a number of practical ways in which they or their institution had sought to overcome the barriers and challenges to the development of DDRS training in mid- and later-career stages. These included: outsourcing training, creating opportunities for collaborative learning and development and enhancing the flexibility of provision (e.g. engaging multidisciplinary perspectives on question-based, challenge-led research topics and opportunities).

2) Strategies and/or Interventions to Implement DDRS Training Vision

Participants identified a number of strategies and interventions that would be needed to achieve buy-in to a new framework from the sector and drive the change necessary to ensure those in the mid- and later-career stages are equipped with the requisite skills to engage with UK and international data infrastructure. These include performance metrics (including REF), criteria for grant funding (particularly at mid- and later-career stages), institutional practices and reward structures. Participants highlighted the need to challenge existing attitudes to training and development and move towards a culture of lifelong learning that enables researchers to develop skills at all career stages in response to the evolving research and innovation landscape, the importance of mentoring for research skills, and the value of a specialist centre or centre of excellence for DDRS training.

3) Brokering a New DDRS Training Framework

Participants were fully in support of a training and development framework that cuts across disciplines and career stages to ensure researchers are able to develop the requisite skills to design, manage, deliver, and disseminate insight from data-driven research. Participants emphasised the importance of a framework that bridges the qualitative and quantitative divide to ensure qualitative DDRS are not overlooked in favour of more 'on trend' (typically quantitative) methods, such as AI and machine learning data analytic strategies. One participant questioned the use of the term 'training framework', however. They argued that a 'professional development framework' would be more appropriate, given skills are gained in a range of formal (training) and informal (e.g. conferences, networks, peer-learning) settings. To ensure the long-term vision is achieved, participants agreed that a skills assessment was needed to identify the range of DDRS (technical and 'soft') required by the sector and the level of knowledge and skills needed by those at different career stages. This would ensure individuals assessed their own skills needs and areas for development within a broader framework designed to address skills gaps and shortages at a sector level.

In reviewing the detailed responses and focus group discussion linked to the core areas summarised above, Steering Group members who attended the workshop noted areas of discussion and feedback that align closely to several areas of DDRS training previously highlighted in their review as possible obstacles and opportunities (see Section 5). Specifically, Steering Group members highlighted areas of workshop recommended priorities, opportunities and challenges:

1) Barriers to Continuing Skills Development throughout a Research Career

- Disciplinary 'silos': Academic careers are still largely tied to disciplinary trajectories, with the result that there is a perceived need among mid- and later-career researchers for discipline-specific training despite the awareness of increased need for interdisciplinary and multidisciplinary collaboration.
- Contextual pressures: Researchers in the social sciences face pressure to publish, with publications clearly linked to authorship, leading to the view that it is better to stick to a special field of expertise, requiring specific methods, rather than innovate or acquire new skills. This is an approach that is, albeit unintentionally, reinforced by the REF.
- HEI priorities: Currently most universities understand skills development to be an individual responsibility at the stages following early-career training and development, with the consequence that there is limited recognition and support given to individuals at later career stages.
- DDRS training Infrastructure beyond doctoral and very early career: While there are training courses available which are not restricted to doctoral and early-career researchers, mid- and later-career researchers are not necessarily aware of this. In addition, it appears common that those seeking training do not have the language in which to communicate their needs and that those providing training do not necessarily describe what they offer in ways that address the needs of researchers. This suggests that demand- and supply-side barriers are interacting to limit the participation of researchers in mid- and late career stages to undertake training.
- Generational factors: Although not all later-career researchers are of the same generation, many will have received their doctoral training before the advent of recent processes of datafication and digitalisation. Some have since acquired digital and data-driven skills in either an informal or formal way, but others have not and are uncertain whether they need to do so and how to do so if they recognise a need. This applies to programming and coding as well as advanced qualitative and quantitative methods.

2) Proposals for How these Barriers Might be Overcome

- Overcoming these barriers will require a shared awareness of the need for and benefits of continuous training in DDRS in HEIs and among individuals. This is important since the aim is to develop an approach or orientation rather than setting specific goals (such as all researchers to be competent in R and

Python). The objective should be the creation of a life course DDRS ecosystem. This requires recognition and reward for continuous skills development throughout the academic/researcher life course. This might take a variety of forms, including:

   a. funding and time recognition in workload allocation at departmental and/or university level. This might include an allowance for researchers at all levels to develop their data-driven research skills and professional competencies. Examples might include attending a training course or workshop, workplace shadowing, participating in a mentoring scheme (as mentor or mentee), committee membership, participating in policy development, public engagement, or knowledge exchange activities;

   b. funding and time recognition for skills training in ESRC and other research grant funding schemes. A change in the way grant and funding applications are framed to enable mid- and later-career researchers to address gaps in their knowledge and skills during the course of the delivery of a proposed project would help to foster a culture of development and build capacity for DDR within the research community. Funding from ESRC for training alone might be considered, but in all cases it should be made clear that what the ESRC is supporting is research excellence – i.e. not skills independent of research questions and agendas;

   c. recognition of outputs other than publications, including software and curated datasets, in promotion and in contributions to REF, as well as recognition of training opportunities in environment statements;

   d. inclusion of evidence of continuous DDRS training as desirable in job descriptions and promotion criteria.

- To recognise the variety of levels of training required, as well as time constraints, and different career trajectories, an array of training formats should be made available. These might include on- and off-line, modular training, problem-based, interdisciplinary and multidisciplinary, peer-learning and collaborative formats. There is demand among later career researchers for both short courses and long-term deep learning over extended periods. A specific need among later career researchers may be to acquire the skills to manage research requiring data-driven research skills.

- ESRC - with UKRI - to develop a shared vocabulary for (social) DDRS. This could support interdisciplinary and multidisciplinary collaboration as well as providing the basis for effective communication.

- A shift in academic culture to address the damaging effects of the ways in which qualitative and quantitative methods are opposed. Specific actions are hard to identify here, but circulation of this report to HEIs, within the ESRC and UKRI, and among professional associations is important, and crucially, inviting responses. In addition, there is a possible role for the Concordat to Support the Career Development of Researchers in bringing about institutional change in relation to promoting life course training. Professional qualifications for specific skills may also be helpful.

- The creation of a platform in which all kinds of DDRS training are located, possibly with an algorithmic matching function, and that includes commercial services.

3) Risks Associated with this Transformative Vision

- There is a possibility that if this vision is implemented in a piecemeal way, the sense of purpose will be lost. This might result in an intensification of existing divides between quantitative and qualitative, conceptual and technical, as well as exacerbating generational divides. The extent to which there will be a buy-in from later career researchers will depend on whether they see their needs being recognised and met. Later career researchers may feel undermined and/or under-valued if the vision is presented in terms of a skills deficit rather than an affirmative response to the need for social science – and world-class social scientists - in a changing environment.

6. Key Findings and Recommendations

The core objective of this scoping study has been to map out the challenges, opportunities and areas of necessary change to existing DDRS training infrastructure in the UK in order to implement a life course model of future social science aligned DDRS training. In this section, we return to the overarching questions for the study and, based on the evidence, summarise the key findings and make recommendations to inform future developments in this area.

Based on a comprehensive Steering Group and stakeholder workshops review and consultation process, the following key findings and recommendations are provided. Aligned to the primary questions outlined at the outset of this scoping study, recommendations are intended to promote and support a life course approach to DDRS training. However, taken together they are designed to create an ecosystem that builds individual capability *and* sector capacity; rather than placing responsibility for training on individuals, they support a vision of an ecosystem for data-driven social science research that is future-proofed.

Key Findings

a) What are Social Scientists' Specific Capability Building and Skills Needs for DDRS Training across the Career Life Course?

- Social science researchers require opportunities to access a life course DDRS training framework that supports acquisition, consolidation and enhancement of qualitative, quantitative and mixed-methods data skills across all stages of career.
- A life course DDRS training infrastructure should be designed to embed a plural model of methods training across all stages of career.
- Life course DDRS training should be designed to be adaptive to new areas of data opportunity, challenge and innovation with training delivery responsive to new computational and digital skills developments, AI and machine learning, coding, programming, digital research methods, visualisation methods and other areas of data innovation.
- Social science researchers would benefit from opportunities to enhance knowledge and training in the areas of data linkage and related strategies to allow maximum utilisation of existing and future data infrastructure resources and opportunities.
- Social science researchers would welcome and benefit from opportunities to move across traditional discipline-based methods silos to allow a more open model of DDRS training to operate across all stages of career.
- Existing CPD models used by practitioners and other professions are a useful comparison and exemplar when considering how to promote continuing education. However, promoting and enabling social science researchers to engage with DDRS at all career stages requires a fundamentally different approach and the development of a bespoke life course DDRS training framework.

b) What Interventions are Required to Enable Social Scientists to Maintain Data-Driven Research across the Career Life Course?

- It is necessary to acknowledge the importance of time, resource availability and access opportunities to promote mid- and late-career engagement with DDRS training opportunities.
- It is necessary to develop formal delivery structures to promote mid- and later-career DDRS training opportunities, recognition and reward mechanisms.
- It is necessary to develop DDRS training incentive opportunities to promote DDRS training engagement across the career life course (e.g., research grant funding aligned incentives).
- It is necessary to develop and support communities of practice that are cross-institutional and cross-sector and accessible across life course career stages.

c) What are the Barriers to Implementing and Supporting a Life Course Model of DDRS Training?

- Access to and improved provision of DDRS training opportunities and infrastructure at mid- and later-career stages represents a fundamental barrier.
- Lack of incentives to promote engagement with and access to DDRS training opportunities at mid- and later-career stages.
- Lack of confidence to declare DDRS training needs at mid- and later-career stages.
- Research organisation (RO) and related impediments to DDRS training opportunities and infrastructure.
- Lack of availability of infrastructure access points across mid- and later-career stages constitutes a significant impediment to life course DDRS training engagement.
- Lack of EDI responsive mechanisms to allow DDRS training engagement across mid- and later-career stages.
- Limited incentive models to move beyond singular or siloed DDRS training modalities to plural and multi-method strategies and fluency.

d) What are the Possible Mechanisms and Strategies to Effectively Design, Deliver and Sustainably Support a Life Course Model of DDRS Training?

- Provide funding mechanisms that encourage life course DDRS training engagement.
- Specify training and development expectations and opportunities through research grants.
- Encourage UKRI-ESRC investments that develop novel DDRS methods and infrastructure to cascade training that is accessible and sustainable.
- Encourage Doctoral Training Partnerships (DTPs) to open up PGR specialist training to mid- and later-career researchers with appropriate funding/support aligned.

- Align RO incentives to promote life course DDRS training engagement and reward (e.g. through REF research environment statements; the Concordat to Support the Career Development of Researchers).

Recommendations

Building on these research findings, a suite of recommendations is provided that spans three proposed implementation themes (1) Implementation Models and infrastructure, (2) Developing a culture/environment for DDRS development across the life course, and (3) Support for researchers.

Theme 1: Implementation Models and infrastructure

- UKRI (including through REF), the HEI sector and research funding bodies such as the Nuffield Foundation, Wellcome Trust and The Leverhulme Foundation should develop and adopt a model and infrastructure to design and deliver training/professional development in DDRS across the life course. This model and infrastructure should:
    - Enable the acquisition, consolidation and enhancement of qualitative, quantitative and mixed-methods data skills across all stages of career.
    - Be adaptive to new areas of data opportunity, challenge and innovation with training delivery responsive to new computational and digital skills developments, AI and machine learning, coding, programming, digital research methods, visualisation methods and other areas of data innovation.
    - Enhance knowledge and training in the areas of data linkage and related strategies to allow maximum utilisation of existing and future data infrastructure resources and opportunities.
    - Articulate the distinction and define the difference between DDRS and more generic research methods training skills needs.
    - Embed a plural model of methods training across all career stages.
    - Promote professional development (rather than 'training' alone)
    - Provide DDRS professional development and related training opportunities that will enable researchers to move across traditional discipline-based methods silos to allow a more multi-lingual model of DDRS training to operate across all stages of career.
    - Invest in and support the creation of an integrated resource platform for information concerning RO and commercial DDRS training opportunities.
    - Establish a DDRS training Centre of Excellence in collaboration with partners (e.g., Turing and/or Ada Lovelace Institute with additional engagement from business, industry, government and wider civil society).

Theme 2: Developing a Culture/Environment for Skills Development across the Life Course

- Funders, Research Organisations (ROs) and other influential stakeholders including national and local government, business and the third sector should collaborate to foster a research culture and environment that promotes the development of DDRS across the life course by providing DDRS training opportunities, incentives, and recognition and reward for mid- and later-career researchers. Such an approach will enable:

  o The alignment of RO incentives to promote and reward life course DDRS training engagement (e.g., through recognition in cases for promotion and merit awards).
  o The establishment of research training funds and access opportunities for mid- and later-career researchers.
  o The specification of training and development expectations and opportunities through research grants.
  o Doctoral Training Partnerships (DTPs) to open up PGR specialist training to mid- and later-career researchers with appropriate funding/support aligned.
  o EDI responsive mechanisms to allow DDRS training engagement across mid- and later-career stages.

Theme 3: Support for Researchers

- To support researchers to upskill in DDR across the career life-course, funders and ROs should:

  o Support the development of communities of practice that are cross-institutional and cross-sector and accessible across life course career stages.
  o Provide funding mechanisms and incentives that encourage life course DDRS training engagement
  o Specify training and development expectations and opportunities through research grants. For example, investments that develop novel DDRS methods and infrastructure could be encouraged to cascade training that is accessible and sustainable.
  o Engage with Research Organisations (ROs) to jointly establish research training funds and access opportunities for mid- and later-career researchers.
  o Link DDRS training opportunities and provision to research environment indicators (e.g., REF requirements) to encourage and reward life course DDRS training investment and future sustained impacts.

It is noted that implementing these recommendations will require further development in a number of areas. In particular, funders and other stakeholders

should collaborate to ensure that the approach taken aligns with and supports the strategic ambitions presented in the Concordat to Support the Career Development of Researchers.

7. Overall Summary

The core objective of this scoping study has been to map out the challenges, opportunities and areas of necessary change to existing DDRS training strategy and infrastructure in the UK in order to implement a life course model of future social science aligned DDRS training.

The scoping study addresses multiple areas of challenge and opportunity in providing recommendations that aim to address this core objective in order to sustainably position the UK as a global leader in the rapidly changing field of DDRS applications and training needs.

Data-driven skills are fundamental to the field of social science research, practice and policy engagement and impact. The type, form, nature, volume and emerging data attributes that social scientists have opportunity to work with are changing rapidly and are likely to continue to evolve and develop at a rapid rate in the coming years.

The UK requires a model of DDRS training that is future-orientated in order to sustainably adapt and prepare for the DDRS needs and opportunities of the present and next generation of social scientists.

This scoping study has implemented a research-led approach to identifying the needs and possible training solutions that will help equip the UK social science community with the requisite data-driven research skills to confidently embrace and maximise opportunities from the complex data attributes that will inevitably populate the data landscapes of the future.

The UK has long been a global leader in the area of data infrastructure development and provision, building on the primary recommendations generated from this scoping study it is proposed that existing DDRS training frameworks should be substantially overhauled if a parallel reputation is to be retained in the area of social science directed DDRS training.

8. References and Further Reading

Baldwin, R. G., & Chang, D. A. (2006). Reinforcing our strategies to support faculty in the middle years of academic life. *Liberal Education, 92(4)*, 28-35.

Baldwin, R., DeZure, D., Shaw, A. & Moretto, K. (2008). Mapping the terrain of mid-career faculty at a research university: Implications for faculty and academic leaders, *Change: The Magazine of Higher Learning*, *40(5)*, 46-55, DOI: 10.3200/CHNG.40.5.46-55

Brannen, J. (2005). Mixing methods: The entry of qualitative and quantitative approaches into the research process. *International Journal of Social Research Methodology*, *8(3)*, 173-184.

Canale, A. M, Herdklotz, C, & Wild, L. (2013). Mid-career faculty support: The middle years of the academic profession. *Faculty Career Development Services,* The Wallace Centre. https://www.rit.edu/provost/sites/rit.edu.provost/files/images/FCDS_Mid-CareerRpt.pdf

Corti, L. & Van den Eynden, V. (2015). Learning to manage and share data: Jump-starting the research methods curriculum. *International Journal of Social Research Methodology*, *18(5)*, 545-559.

Couper, M. (2013). Is the sky falling? New technology, changing media, and the future of surveys. *Survey Research Methods*, *7(3)*,145-156.

Crow, G. (2020) Hedgehogs, foxes and other embodiments of academics' research career trajectories. *Contemporary Social Science, 15(5)*, 577-594.

Degn, L., Franssen, T., Sørensen, M. P, de Rijcke, S. (2018). Research groups as communities of practice – A case study of four high performing research groups. *Higher Education, 76*, 231-246.

Department for Business, Energy and Industrial Strategy. (2021). *Research and development (R&D) people and culture strategy*. https://www.gov.uk/government/publications/research-and-development-rd-people-and-culture-strategy

Department for Business, Energy and Industrial Strategy and Department for Digital, Culture, Media and Sport. (2018). *Artificial intelligence sector deal*. https://www.gov.uk/government/publications/artificial-intelligence-sector-deal#full-publication-update-history

Durrant, G. & Lang, I. (2004). *Consultation exercise report.* Discussion Paper. Unpublished. National Centre for Research Methods, University of Southampton. https://eprints.ncrm.ac.uk/id/eprint/114/

Durrant, G., Luff, R., Wiles, R. & Crow, G. (2015). *Consultation on training needs in advanced social science research methods.* National Centre for Research Methods, University of Southampton. https://eprints.ncrm.ac.uk/id/eprint/3765/

Economic and Social Research Council. (2022). *ESRC strategic delivery plan 2022 – 2025.* https://www.ukri.org/publications/esrc-strategic-delivery-plan/

Economic and Social Research Council. (2022*). ESRC data infrastructure strategy 2022 – 2027.* https://www.ukri.org/publications/esrc-data-infrastructure-strategy/

Economic and Social Research Council. (2022). *ESRC postgraduate training and development guidelines 2022.* https://www.ukri.org/publications/esrc-postgraduate-training-and-development-guidelines/

Economic and Social Research Council. (2018). *Knowledge and skills needs call for evidence: Summary report and next steps.* https://webarchive.nationalarchives.gov.uk/ukgwa/20200930152854/https://esrc.ukri.org/news-events-and-publications/news/news-items/knowledge-and-skills-needs-call-for-evidence-summary-report-and-next-steps/

Ferretti, A., Ienca, M., Velarde, M.R., Hurst, S. & Vayena, E. (2022). The challenges of big data for research ethics committees: A qualitative Swiss study. *Journal of Empirical Research on Human Research Ethics*, *17(1-2)*, 29-143.

Ferrie, J., Wain, M., Gallacher, S., Brown, E., Allinson, R., Kolarz, p., MacInnes, J., Sutinen, L., & Cimatti, R. (2022). *Scoping the skills needs in the social sciences to support data-driven research.* UK Research and Innovation. https://www.ukri.org/publications/scoping-the-skills-needs-in-the-social-sciences-to-support-data-driven-research/

Hollywood, A., McCarthy, D., Spencely, C. & Winstone, N. (2019). Overwhelmed at first: The experience of career development in early career academics. *Journal of Further and Higher Education*, *44(7)*, 998-1012.

Kilburn, D., Nind, M. & Wiles, R., (2014). *Short courses in advanced research methods: Challenges and opportunities for teaching and learning.* National Centre for Research Methods, University of Southampton. https://eprints.ncrm.ac.uk/id/eprint/3601 Accessed 23.11.2022

Kim, J. (2015). Competency-based curriculum: An effective approach to digital curation education. *Journal of education for library and information science*, *56(4)*, 283-297.

Lave, J. & Wenger, E. (1991). *Situated learning: legitimate peripheral participation*. Cambridge University Press.

Leon, S. & Brennan, S.A. (2015). *Scholars as students: Introductory digital history training for mid-career Historians.* Roy Rosenzweig Center for History and New Media, George Mason University.

http://mars.gmu.edu/bitstream/handle/1920/9825/RRCHNM_DoingDH_WhitePaper_8-31-2015.pdf?sequence=1&isAllowed=y

Lewthwaite, S. & Jamieson, L (2019). Big-qual - Why we should be thinking about qualitative data for research, teaching and policy. *LSE Impact of Social Science.* https://blogs.lse.ac.uk/impactofsocialsciences/2019/03/04/big-qual-why-we-should-be-thinking-big-about-qualitative-data-for-research-teaching-and-policy/

Luhmann, J. & Burghardt, M. (2022). Digital humanities—A discipline in its own right? An analysis of the role and position of digital humanities in the academic landscape. *Journal of the Association for Information Science and Technology, 73(2)*,148-171.

Lukianova, L. (2016) Motivation Factors of Adult Learning. *The New Educational Review, 44(2)*, 223-229.

Mantai, L. (2019). @Feeling more academic now": Doctoral stories of becoming an academic, *The Australian Educational Researcher, 46*, 137-153.

Marres, N. (2017) *Digital Sociology: The reinvention of social research*. Polity.

McAlpine, L., Amundsen, C. &Turner, G. (2014). Identity-trajectory: reframing early career academic experience. *British Educational Research Journal, 40(6)*, 952–969.

Millard, D. (2016). Reconciling work and family in the context of boundaryless careers: Academic careers in social sciences and humanities. *Ex Aequo, 34*, 111-126.

Muniyappa, K. (2007). The role of postdoctoral training for careers in research and higher education. *Current Science*, 92(4), 450–454.

Nind, M., Holmes, M., Insenga, M., Lewthwaite, S., Sutton, C. (2019). Student perspectives on learning research methods in the social sciences. *Teaching in Higher Education, 25(7)*, 797-811. https://doi.org/10.1080/13562517.2019.1592150

Office for Artificial Intelligence, Department for Digital, Culture, Media & Sport, and Department for Business, Energy & Industrial Strategy. (2021). *National AI Strategy*. https://www.gov.uk/government/publications/national-ai-strategy

Owen-Pugh, V. (2007). Theorizing sport as a community of practice: The coach–athlete relationship in British professional basketball. In Hughes, J., Jewson, N. & Unwin, L. (Eds.) *Communities of practice: critical perspectives* (pp.189-215). Routledge.

Rogers, R. (2019). *Doing Digital Methods*. Sage.

Rosemberg, C., Allinson, R., De Scalzi, M., Krčál, A., Bryan, B., Farla, K., Dobson, C., Cimatti, R., Wain, M., & Jávorka, Z. (2022). *Evaluation of the Q-Step programme*.

Nuffield Foundation. https://www.nuffieldfoundation.org/publications/q-step-evaluation

Tazzyman, S., Moreton, R., Bowes, L., Wakeling, P. and Stutz, A. (2021). *Review of the PhD in the Social Sciences*. Economic and Social Research Council. https://www.ukri.org/wp-content/uploads/2022/03/ESRC-020322-Review-of-the-PhD-in-the-Social-Sciences.pdf

UK AI Council. (2021). *AI Roadmap*. https://www.gov.uk/government/publications/ai-roadmap

UK Research and Innovation. (2022). *Transforming our world with AI*. https://www.ukri.org/about-us/strategy-plans-and-data/strategies-and-reviews/ai-review-transforming-our-world-with-ai/

Unwin, L. (2007) English apprenticeship from past to present: The challenges and consequences of rampant 'community' diversity. In Hughes, J., Jewson, N. & Unwin, L. (Eds.) *Communities of practice: critical perspectives* (pp.109-119). Routledge.

Wenger, E. (1998). *Communities of practice: learning, meaning, and identity*. Cambridge University Press.

Wenger, E., McDermott, R. & Snyder, W. (2002) *Cultivating Communities of Practice: A Guide to Managing Knowledge.* Harvard Business School Press.

Whitchurch, C. & Gordon, G. (Eds.). (2009). *Academic and Professional Identities in Higher Education: The Challenges of a Diversifying Workforce*. Taylor & Francis Group.

Wiles, R., Durrant, G., De Broe, S., & Powell, J. (2005_. *Assessment of Needs for Training in Research Methods in the UK Social Science Community*. Discussion Paper. Unpublished. https://eprints.ncrm.ac.uk/id/eprint/4468/

Wiles, R., Durrant, G., De Broe, S. and Powell, J. (2009). Methodological approaches at PhD and skills sought for research posts in academia: a mismatch? *International Journal of Social Research Methodology*, *12(3)*, 257-269.

Appendix 1: Steering Group Membership

Professor Gordon Harold (Chair; University of Cambridge)

Professor Celia Lury (University of Warwick)

Professor John Goodwin (University of Leicester)

Dr. Sarah Lewthwaite (University of Southampton)

Professor Jackie Carter (The University of Manchester)

Additional contributors:

- Dr. Laurie Parsons (University of Leicester)
- Dr. Helen Watts (University of Cambridge)
- Mr. Gregory Campbell (The University of Manchester)