SHAPING DATA AND SOFTWARE POLICY IN THE ARTS AND HUMANITIES RESEARCH COMMUNITY

Policy Recommendations for the AHRC

Rebecca Taylor, Johanna Walker, Simon Hettrick, Philippa Broadbent (University of Southampton, Software Sustainability Institute), and David de Roure (University of Oxford)

July 2022
## CONTENTS

**5.1 Changing culture**
- 5.1.1 Implementing a specific data policy for the AHRC 3
- 5.1.2 Clear support for data, software and digital skills 3
- 5.1.3 Move towards data and software sharing 4
- 5.1.4 Legitimise support for research technical professionals 4
- 5.1.5 Shared guidance across research councils 4

**5.2 Improved training in data, software, and digital skills**
- 5.2.1 Building an understanding of core digital skills 5
- 5.2.2 Embed digital skills in doctoral training 5
- 5.2.3 Signpost training and resources 6
- 5.2.4 Build communities of research practice 6

**5.3 Research funding**
- 5.3.1 Strategic interdisciplinary programmes 8
- 5.3.2. New fellowships 8
The study identified five broad areas over which changes to the AHRC policy, and well directed interventions, could shift the arts and humanities research community towards better practices with data and software, and strengthen digital skills.

The follow recommendations are made across three topics:

1. **Changing Culture**
   - Rather than looking for significant investment, this topic focuses on changing research culture around data, software and digital skills.
   - Total investment over five years: £160,000.

2. **Improved training in data, software and digital skills**
   - This topic includes activities to understand skills requirements, embedding skills into doctoral training, signposting available training and developing self-sustaining communities.
   - Total investment over five years: £9.5m.

3. **Research funding**
   - Focuses on two new opportunities around strategic interdisciplinary programmes and fellowships.
   - Total investment over five years: £28.0m.

**5.1 CHANGING CULTURE**

Academic culture drives the practices and techniques that are used in research. By promoting how it values data, software and digital skills, the AHRC has a considerable amount of influence over the direction of research culture. The following policy suggestions do not have associated external costs (outside of its staff time), apart from those associated with further research, but they could have a considerable impact on improving the use of data and software, and hence a transformational role in advancing research across the community.

**5.1.1 Implementing a specific data policy for the AHRC**

The data policy used by the AHRC simply refers to the generic UKRI policy. This has two clear disadvantages. First, it means that the policy is not specific to the needs of the arts and humanities research community. Second, and possibly far more importantly, a lack of a specific policy intimates that data is not important to the AHRC.

The AHRC should review the data policies of the UKRI and the other research councils and develop its own policy on data and software, addressing the diversity and covering the specific needs of the arts and humanities research community. It should also relate to data and software practices at all stages of the research process, rather than focusing specifically on open data. To raise awareness, the new policy should be displayed prominently on the AHRC's website and included in all announcements of funding opportunities. That policy would provide the foundations for clear guidance supporting data and software practices and skills.

**5.1.2 Clear support for data, software and digital skills**

Drawing on the AHRC’s data and software policy, the vital role of data, software and digital skills should be made explicit in AHRC research funding guidance that applies across its funding opportunities. This information should be readily available on the AHRC website, included in all its funding opportunity documents (potentially in abbreviated form), and should be linked to all other documentation that supports the funding process. This would clearly establish the AHRC’s support for data, software and digital skills across all research.

An important aspect of raising the profile of data, software and digital skills is that peer review panelists must be briefed on this new policy. No policy change can have effect until everyone who applies the policy understands that the landscape has changed. Panel guidance must be updated. Not all experienced panelists will regularly review the guidance, so all panelists must be briefed about the change in policy and the seriousness with which the AHRC intends to apply it.

To build evidence of the role of software and data in research, all funding opportunities should include a description of data- and software-focused dissemination strategies, for example, whether the AHRC expects data and software to be registered with ResearchFish, and under what circumstances data and software should be cited in presentations and publications.
5.1.3 Move towards data and software sharing

The development of a policy on data and software (recommendation 5.1.1) is a step towards the creation of a culture of good practice in data and software use within the arts and humanities. The next step is the promotion of open practices, e.g., licensing, DOIs, repositories, FAIR data and software.

The AHRC should take the position that data and software generated during any AHRC-funded project should be shared, unless there is a rationale for it not to be. As identified by this study, there is work to be done to convince the community about the benefits of openness, the AHRC should focus on describing and evidencing these benefits. It would also be necessary to discuss any drawbacks and problems caused by openness, which can be mitigated by weighing these issues against the benefits of open research.

The AHRC should fund a study of open research in the arts and humanities research community, which should include a review of existing literature, a comparison of policies and practices adopted by other councils, and talking with stakeholders and groups within the community about their experiences and concerns with sharing, exploring the adequacy of infrastructure such as repositories and archives. The conclusion should be the construction of best practice case studies to promote the benefits and explore solutions where there are concerns. These case studies should be used to increase awareness of the benefits of openness for the research community and enable campaigns to improve infrastructure where this is a barrier to sharing. Ultimately, the study would prepare the ground for a move towards most AHRC-funded data and software being open licensed.

COST

Developing a methodology to study open research in the arts and humanities research community, conducting the study, which would require resources from two researchers working at 0.5 FTE with oversight from a principal investigator at 0.2 FTE. The study would require a year to adequately develop and then deploy a methodology, produce case studies and disseminate the findings, indicating a cost in the region of £160,000.

Five year total: £160,000.

5.1.4 Legitimise support for research technical professionals

Access to research technical professionals, such as Research Software Engineers and data stewards, is vital if data and software are to be reliably developed, managed and archived. The AHRC guidance must clearly state that these professionals are legitimate and, therefore, endorsed positions on bids for its funding opportunities.

5.1.5 Shared guidance across research councils

The issues that surround the use of data and software in research are one of the few areas that apply equally across all disciplines. All of the research councils are working to address these issues, and this presents a significant opportunity for the AHRC to benefit from the experience gained across the UKRI. The AHRC should continue to strengthen its work with other research councils to share knowledge and expertise and, where appropriate, enter into joint funding of activities that benefit data, software and digital skills. The UKRI’s Digital Research Infrastructure\(^1\) (DRI) programme may be an obvious target for this cross-council activity.

---

1 Digital research infrastructure – UKRI
5.2 IMPROVED TRAINING IN DATA, SOFTWARE, AND DIGITAL SKILLS

Increasing access to digital skills requires investment in training and improved access to training. Training should focus on providing a broad foundation of skills across the community from basic data literacy, data management, data sharing, research ethics, data manipulation techniques and software engineering.

5.2.1 Building an understanding of core digital skills

When it comes to establishing a ‘baseline of digital skills’, it must be understood that there is no universal set of core digital skills that will apply across the entire arts and humanities research community. There will be much overlap in the digital skills necessary within each discipline and area, but there will also be variations.

To provide some guidelines on what a baseline of digital skills might look like, the AHRC should fund a study to investigate the most commonly required skills across different parts of the arts and humanities research community. This work should encompass staff working in academia, IROs, the GLAM sector, as well as looking at the private sector to gain an understanding of the skills required by employers. The AHRC’s training grant scheme partners and other stakeholders should be consulted in order to understand how this baseline might be delivered within training to equip graduates with the appropriate subset of the core digital skills (recommendation 5.2.2).

Skills requirements change as digital techniques evolve, so this snapshot study of digital skills would need to be refreshed, with a cadence of every five years likely to gain the optimum balance between being cost effective and up to date. Oversight of this work to understand digital skills could be provided by the Digital Skills Hub (recommendation 5.2.3) once it is operational.

**COST**

Developing a methodology to study digital skills; conducting the study, which would require resources from two researchers working at 0.5 FTE with oversight from a principal investigator at 0.2 FTE. The first study would require a year to adequately develop and then deploy a methodology, indicating a cost in the region of £160,000. Further studies would reuse the methodology and are expected to require the same number of staff, but run for only six months, indicating a cost of £80,000.

**Five year total:** £160,000.

5.2.2 Embed digital skills in doctoral training

AHRC doctoral training grant schemes are effective vehicles for supporting digital skills training in the arts and humanities research community. All doctoral students of the future should be expected to have or acquire a baseline of research skills (drawn up in the research conducted in 5.2.1). This training would be offered by training grant partners and take place in students’ first year of study, to allow them to underpin their research with digital skills from the start of their doctoral journey. Flexibility is key to this requirement: whilst the baseline training would be required, the type and depth of training may change depending on the field of study. Doctoral training partners would have some flexibility to shape what is offered, and some students may be completely exempt from learning digital skills, but this should be the exception and not the rule.

In addition to baseline training, AHRC training grant schemes would provide all doctoral students with access to a personal development fund that they can use flexibly to learn new research tools and techniques and innovate in their practice. The AHRC must be clear in communicating to students and supervisors that this allocation can be spent on acquiring a broad range of skills and knowledge. Materials and courses could be endorsed by the AHRC via the new Digital Skills Hub (see 5.2.3, below).

**COST**

It is noted that making a personal development budget available to PhD students is only part of the solution to training. To be effective, students must be given the power to choose how to spend their budget. A development budget of a minimum of £3,000 per student over the duration of their PhD would significantly increase the training that students could obtain.

The AHRC currently funds 10 DTPs. Assuming 100 students per DTP, and including 40 CDP studentships, suggests a budget of £3.1m would be required in total, or approximately £0.78m a year (assuming a
Five year total: £3.9m.

5.2.3 Signpost training and resources

A significant barrier to acquiring digital skills is caused by the difficulties that AHRC researchers report in identifying suitable training of an appropriate duration, level and accessibility. The AHRC can help by promoting suitable training courses, run by their own funded partners, similar to the approach that the ESRC has taken with the National Centre for Research Methods\(^2\) (NCRM). In addition, they can promote training provided by other UKRI funded partners such as the Turing Institute, the Software Sustainability Institute, and IROs such as the British Library or the National Archives. Finally, they might want to promote training undertaken and rated by researchers in the community - a living list of training, in which courses are suggested, tried out and either stay on the list or are removed based on feedback from researchers. Other important digital resources could also be included, such as details about RSE groups and other skilled professionals who can help researchers. There are many approaches to this problem, all of which would require long-term support for at least five years in the first phase to successfully build a user community around the resource - a Digital Skills Hub.

The Digital Skills Hub would be overseen by a director and a project manager. Four digital training coordinators would collect training links from the community and build networks with training providers. A communications officer would raise awareness for the hub and undertake promotion with research groups and projects via press and social media. These seven staff are the minimum required to make the hub feasible.

The hub would present information about training and allow it to be filtered to identify, for example, specific types of training, training within a cost range, and whether it is in-person or online. A review system would allow researchers who have completed training to provide feedback on the course. The hub could also oversee other elements of the recommendations including the embedding of skills in doctoral training (recommendation 5.2.1) and building of communities of practice (recommendation 5.2.4).

The Digital Skills Hub team could be created and based within the AHRC or it could be outsourced through a funding opportunity. If the hub is to be the focus of a funding opportunity, it is vital that the winning bid must include plans for packaging and handing over to the AHRC or to a new partner. In this way, its sustainability would be assured if the first winner of the bid can no longer run the project.

COST

The seven staff members would cost around £880,000 a year.

Website development and hosting would require around £15,000, with a further £5,000 per year of maintenance from a consulting web developer.

A significant amount of UK travel would be required, around 6 trips per year for the director and each digital training coordinator, which necessitates a travel budget of around £15,000 per year.

Five year total: £4.6m.

5.2.4 Build communities of research practice

The AHRC could directly address the lack of a forum to discuss data, software and digital skills within the arts and humanities research community by providing funding for conference papers that showcase the latest data, software and digital developments within the community. There are two strands to this recommendation:

5.2.4.1 Domain conferences:

Researchers who have expertise in data, software and digital methodologies have reported difficulties in finding funding to present their research at traditional arts and humanities domain focused conferences and events. This prevents these researchers from showcasing the benefits of data, software and digital skills in the context of a disciplinary knowledge exchange. To overcome this problem, a conference fund should be made available to researchers that would pay for the conference fee travel and accommodation. Funding could be provided as a standing opportunity twice a year to which researchers can apply. They would need to name the conference and provide details of their accepted paper or workshop, explaining how it would contribute to knowledge around data and software.

---

\(^2\) ESRC National Centre for Research Methods (ncrm.ac.uk)
strategies and tools.

Another mechanism would be to channel funding to specific conferences, which would then dispense conference registration and travel bursaries for researchers whose abstract submission meets the requirements to disseminate data and software practice. Not only would this reduce the management overhead on the AHRC, it would also be an important signpost that the AHRC values data, software and digital skills.

**COST**

Assuming that 50 researchers will require this funding each year, an average registration fee of £200 and an average one-to-two-day travel budget of £800, suggests an annual budget of £50,000. These funds could be managed and dispensed by the Digital skills Hub.

**Five year total: £250,000.**

5.2.4.2 Digital data and software conferences:

This would enable researchers to share knowledge, network and build collaborations but, unlike traditional conferences, a focus on data and software practices and tools would be actively encouraged. The conference could replicate the model of the ESRC’s NCRM Summer School. There is significant interest in digital research skills within industry, so it is expected that a conference of this type could attract significant sponsorship after its first year, and this would provide a path to the conference to become self-sustaining. The creation of a novel, popular and self-sustaining conference would be a significant achievement for the AHRC’s digital strategy.

The skills conference should be planned to take place each year for a period of at least five years. A series of conferences will help to build a community, which will grow faster and have significantly more cumulative impact than the conferences alone.

**COST**

Assuming 150 attendees at a mid-priced venue, with initially a nominal entry fee and a travel bursary of £200 for 50 of the attendees (to ensure attendance by early career researchers), suggests a conference budget of £43,000 per year should be sufficient. Conference sponsorship of around £10,000 per year should be relatively straightforward to identify within the first five years. Following this period, the conference should have reached sufficient popularity to allow a more substantial registration fee to be charged (around £200 per person would ensure the conference breaks even).

The organisation of the conference would require 9 months’ effort of a staff member, leading to a staffing cost of £45,000. Organising the conference could fall under the remit of the Digital Skills Hub, which would remove the need to hire additional staff and remove this cost.

**Five year total: £600,000 (reduced to £160,000 if sponsorship is gained and staff from the Digital Skills Hub organise the conference).**
5.3 RESEARCH FUNDING

Research funding is the area in which the AHRC has the greatest level of control over the research community. The following investments would build digital skills in the AHRC community. However, this series of investments must be accompanied by the policies described in section 5.1. Investment alone will have considerably less impact than a combination of investment and the implementation of policies to change culture within the arts and humanities.

5.3.1 Strategic interdisciplinary programmes

Some groups within the arts and humanities research community show significantly higher levels of expertise with data and software. AHRC investment into data and software is needed, but this separation of skills across groups means that investment with a purely focus on digital techniques is likely to preclude less data-and-software experienced researchers from applying, although it is this group that would most benefit from participating in these projects. To address this issue, and to increase knowledge exchange across the community, the AHRC should tailor interdisciplinary funding opportunities that promote cross-fertilisation of skills across groups within the AHRC or, more broadly, with groups outside of the AHRC remit.

The 2022 AHRC funding opportunity “Embed digital skills in arts and humanities research” provides a potentially useful way to encourage this type of dissemination of skills within the community. Successful applicants will run digital training aimed at researchers in the community. If it is to be successful, the reviewers and review panel need to reflect the desired composition of the applicants. It will be important to evaluate both the take up of this skills training when the programme is rolled out and the outcomes for providers and researchers, to understand the reach of the training within the community.

Another mechanism is to design funding opportunities with an embedded requirement for interdisciplinarity. This might mean forging links between digital humanities and non-digital researchers within the arts and humanities, or it might mean linking arts and humanities researchers with researchers in other disciplines such as computer science, physics, biology or social sciences. The goal is to encourage interdisciplinary working, as a way to ensure that the cross-fertilisation of skills within the research team is emphasised and facilitated by the structure of the funding.

COST

Based on the existing skills call, £1.5m per year.

Five year cost: £7.5m.

5.3.2 New fellowships

Fellowship opportunities have the power to significantly increase interest in a target field. Not only do they provide long-term funding for experts to showcase their work in the field, but the competition for positions raises awareness of the field across a much broader community than just those who are successful in their bids. A good example of this effect is provided by the Research Software Engineer Fellowships run by the EPSRC and later the STFC. These provided funding for around 3-5 fellowships each year but, with hundreds of intents to submit, raised awareness of research software engineering across most research universities.

The AHRC should either run its own research software engineering fellowship scheme or potentially join forces with the EPSRC and STFC’s funding opportunity. The fellowships should last five years to allow fellows the opportunity to build a network to disseminate their skills. There is much opportunity to broaden from research software engineers to research technical professionals in general, which would provide the AHRC with an opportunity to be the first funder to run such a funding opportunity. Fellows could be generalists working across fields, or they could work within a particular domain and focus on disseminating domain-specific data, software and digital skills to that domain.

An RSE or RTP funding opportunity could be further enhanced by providing funding for a junior position alongside the fellow. This person could be trained during the fellowship to double the impact of the call and start to build an understanding of how to develop new Research Software Engineer and Research Technical Professional roles in the AHRC.

---

3 Embed digital skills in arts and humanities research – UKRI
4 Research Software Engineer fellowships 2020 – UKRI
**COST**

Based on the EPSRC and STFC’s successful fellowships, each five-year fellowship would require an average of £0.8M. Around 5 fellowships would be required each year to ensure that applicants felt there was a significant enough chance of success to risk taking the time to apply. Hence 25 fellowships would be required in total.

**Five year cost: £20m.**