iDAH Research Software Engineering (RSE) Steering Group Working Paper

Table of Contents

1. Background

This working paper resulted from two (hybrid) workshops conducted in May and June 2022, chaired by Professor James Smithies (King's College London) at the request of Tao-Tao Chang, AHRC Head of (Research) Infrastructure. The workshops were conceived and organised by Dr. Anna-Marie Sichani (AHRC Policy and Engagement Fellow), and hosted at The Alan Turing Institute. Contributors are listed in Appendix 4.

2. Purpose and Audience

Our **purpose** was to convene a broad and representative group from the UK Research Software Engineering (RSE) community to discuss opportunities for and barriers to the development of Arts & Humanities (AH) RSE capability, with a specific focus on contributing to the AHRC Infrastructure for Digital Innovation and Curation in Arts and Humanities (iDAH) project and a wider remit to consider longer term strategic priorities and opportunities for alignment with UKRI and EU initiatives. The discussion was intended to be foundational, inclusive, and broad-ranging, involving a wide stakeholder group encouraged to engage in 'blue-sky' thinking over short, medium, and long-term time horizons. The analysis contained in this working paper should be read in that context, as a reflection of early stage discussions intended to provide a platform for future more focused activity. Additional discussion and analysis is needed to produce substantive actionable conclusions.

Our primary **audience** is AHRC. The working paper will be shared with appropriate committees within AHRC and UKRI, relevant stakeholder groups such as the Society of RSE (SocRSE) and Software Sustainability Institute (SSI), but not be made publicly available due to its preliminary nature.

3. Professional Identity and Policy Development

The workshops documented in this paper represent the broadest (although by no means the first) conversation about the UK AH RSE community undertaken. This is largely a function of growth in awareness of the field over the past five years, enabled by initiatives such as The Society of Research Software Engineering (SocRSE) and The Software Sustainability Institute (SSI). These initiatives have raised awareness of a category of Research Technology Professionals (RTP) that work at the intersection of research and information technology, primarily in universities and cultural institutions, inclusive of RSEs. Their primary

expertise often relates to programming and tool development designed to answer research questions, store and make available digital content, and publish research findings but RSEs are involved in a very wide variety of activities ranging from project leadership/management and research / business analysis to UI/UX design, systems administration, and collections management. They also often lead or contribute to systems and infrastructure development and maintenance in partnership with both internal IT departments and external vendors. They are an increasingly key component in interdisciplinary, computationally intensive research, working alone and in teams and across an extraordinary range of research disciplines. It is important to recognise that many people who could identify as RSEs might either not choose to do so, or not be aware of the RSE community. Technical research activity and infrastructural development and maintenance has been occurring for decades, and field building is ongoing. Because their skills are in high demand in the commercial and government sector, competition for RSEs is intense and there is broad acknowledgement that investment needs to be made to ensure adequate training, remuneration, and career pathways are available.

AH RSEs are therefore a subset of a much larger cross-disciplinary and cross-sector group of RSEs who "combine expertise in programming with an intricate understanding of research" and research infrastructures.[1] This provides an important commonality of professional culture and purpose, alignment to existing and evolving professional practices and policies, and enables them to facilitate knowledge transfer across sectors and knowledge domains. They are an invisible workforce in some ways, working in enabling roles at the intersection of research and technology, but have become central to the strategic goals of UKRI and AHRC, and will be essential to the success of initiatives such as iDAH, Towards a National Collection (TaNC), and future Strategic Priorities Fund (SPF) investments. It is essential that their expertise is captured and contributes to AHRC strategic planning (in technical and non-technical contexts) over the coming decades.

In addition to being part of a wider professional community, where the challenges of working in interdisciplinary groups bring valuable new skills and methods from other fields, RSEs have significant domain knowledge gained through domain-specific work within their immediate stakeholder community. In many (but not all) cases, AH domain knowledge is essential to the delivery of high quality research and impact outcomes. This requires attention to domain-specific AH RSE experiences and requirements, to maximise the

---

[1] Software Sustainability Institute (SSI), 'Research Software Engineers', *Software Sustainability Institute*, https://www.software.ac.uk/research-software-engineers. Accessed 07 July 2022.

potential for RSEs to contribute to key activities and investments across AHRC stakeholder communities. Policy development needs to occur at this domain-specific level, enhancing RSE support for the AH research community (including the Digital and Computational Humanities) and building the capacity for AH RSE to contribute their valuable perspectives to the development of the national RSE community, as well as the national AH community. This activity is more pressing because dedicated or experienced AH RSE skills are rare, with pockets of capability in a handful of institutions and funding models that currently offer advantages to institutions capable of and motivated to provide internal funds to achieve their research goals. The relationship between AH RSE and the broader Research Technology Professional (RTP) community and associated initiatives such as the Technician Commitment can be expected to be enhanced through such activity.

4. Policy Context and Alignment

Over the past decade the role of software development is increasingly recognised by the UK research community and the contribution of Research Software Engineers (RSE) to UK research and science has grown markedly. The RSE community began in March 2012 at the Collaborations Workshop, in Oxford, when a small group met to discuss the lack of recognition, careers and, more importantly, a name for software developers in academia. That meeting would eventually lead to a nationwide awareness raising campaign, a vibrant international community, and a new role in research: the Research Software Engineer.[2]

In 2013 the UK Research Software Engineers Association was established as the first professional body to represent the RSE community. In March 2019 the Society of Research Software Engineers[3] was launched as a charitable incorporated organisation, replacing the previous body. One of the most significant developments has been the creation of several RSE groups in research institutions, the first of which was set up at UCL in 2012. Since then over 25 RSE groups have been set up across the UK, and a growing number internationally. In order to keep up with the phenomenal growth of the RSE community a Research Software Group Leaders Network[4] was established in 2015 and a bi-annual survey[5] run across all

---

[2] https://www.software.ac.uk/blog/2016-08-17-not-so-brief-history-research-software-engineers-0?_ga=2.116467334.1469987827.1658848569-448543463.1658848569

[3] https://society-rse.org

[4] https://software.ac.uk/blog/2015-08-07-new-network-support-research-software-groups-across-uk?_ga=2.156104690.891793815.1561539411-2136342297.1555488175

[5] https://www.software.ac.uk/news/complete-international-rse-survey-and-help-us-understand-rse-community

the national RSE associations[6] in order to capture insight into the community in terms of their demographics, impact and careers and on how best to support it from funders' and other policymakers' perspective. King's Digital Lab (KDL) at King's College London was one of the first Arts & Humanities teams to start using the designation RSE, using the term internally from 2018 and publishing their role descriptions in 2019.[7]

Over the years there have been several initiatives and reports providing further evidence-based context around the value and the place of the RSE community in UK academia. The State of the Nation report (Brett et al 2017), published in 2017, described the growth of the RSE community and further established the requirements around RSE roles, such as training provision, recognition and career paths. The recent Research Software Engineer Knowledge Integration Landscape Review 2021, as part of the ExCALIBUR Project (Parsons et al 2021), offers a comprehensive review of the skills required by Research Software Engineers in High Performance Computing (HPC) and their future training needs. These reports not only map the RSE community as it evolves but situate it within the broader funding and policy context of Open Research and UK Research and Innovation.

The RSE community has also been recognised and supported by funding bodies for its contribution: since 2015, the Engineering and Physical Sciences Research Council (EPSRC) has been releasing a fellowship call (2015, 2017, 2020) specifically targeted at Research Software Engineers, to support the role of RSEs and establish a cohort in the academic environment[8]. It is rare, however, to find adequate provision for RSE funding and support outside STEM, especially for Arts and Humanities, for instance the 2020 round specifically excluded AH engagements.

A number of institutions have been also supporting and championing the availability and integration of software skills in UK research and the RSE community more broadly. Since 2010, the Software Sustainability Institute (SSI)[9] has facilitated the advancement of software in research by cultivating better, more sustainable, research software to enable world-class research ("Better software, better research"). Over the years, the Institute has advocated, alongside others, for the fundamental importance of research software for UK research, the recognition of software as a research output, and the importance of reproducible research. Through its well-known Fellowship Programme, surveys, RSE conferences and a number of

---

[6] https://researchsoftware.org/2021/01/27/introducing-the-international-council-of-RSE-associations.html

[7] Smithies, James. Research Software (RS) Careers: Generic Learnings from King's Digital Lab, King's College London. Zenodo, February 7, 2019. Accessed March 29, 2019. https://zenodo.org/record/2564790.

[8] https://www.ukri.org/opportunity/research-software-engineer-fellowships-2020/

[9] https://www.software.ac.uk

relevant activities and events, SSI actively nurtures and supports a cross-disciplinary network of RSEs advocating for research software best practices.

The Alan Turing Institute[10] has made significant contributions to field development, strategically placing Research Software Engineering at the centre of its work and organisational mission, with the Research Engineering team[11] contributing RSE and Research Data Scientist skills, collaboratively delivering impactful research across the Turing's programmes. The Turing was one of the first UK institutions to cultivate and promote an AH RSE community, by bringing together Humanities with Data Science and Research Software Engineering through discussion groups, meetings, workshops, and joint research projects in the area of digital humanities and by showcasing how research software engineering and data science can answer questions relevant to the arts and humanities through an interdisciplinary ethos. The Turing Humanities & Data Science Interest Group (H&DS) focuses on leveraging the profile of data-driven humanities research at Turing and encouraging collaborations among other UK stakeholders such as the British Library, the National Records of Scotland and The National Archives in the UK. This group often organises digital humanities data study groups and has published a Data Science and digital humanities white paper, where a plea for RSE provision for digital humanities has been made (McGillivray, Barbara et al., 2020) . Most recently, the Turing hosted the first Digital Humanities & Research Software Engineering virtual summer school (26 - 30 July 2021) and has an ambitious programme of catalyst AH RSE activities, Accelerating AI in the Arts & Humanities.

5. Analysis and Key Recommendations

The AH RSE community has matured significantly in the last five years, benefiting from national RSE initiatives outlined above and commitment to interdisciplinarity and cultural change across UKRI. It has also become apparent to key stakeholders that historic investments in Digital Arts & Humanities spanning decades and including the Arts & Humanities Data Service (AHDS) suffered from a lack of investment in human infrastructure capable of improving sustainability and impact. Greater attention to cybersecurity, and practices such as research data management (RDM), has also raised the profile of AH RSEs and highlighted inconsistencies in funding, capacity, and career pathways. High profile investments such as Living with Machines[12] (who benefitted from an AH RSE as acting PI) and Towards a National Collection have demonstrated the pressing need to improve AH

---

[10] https://www.turing.ac.uk/
[11] https://www.turing.ac.uk/research/research-engineering
[12] https://livingwithmachines.ac.uk/

RSE capacity, and tap into the collective wisdom of the AH RSE community. Major infrastructural initiatives in Europe, including long-standing projects such as DARIAH and CLARIN, and the recent European Cultural Heritage Cloud, have also highlighted the central importance of RSEs to the success of major strategic initiatives. Recent discussions about the formation of a UK - Ireland Digital Humanities association have emphasised the importance of RSEs to that initiative too.

These changes intersect with the technological convergence experienced across the entirety of the UKRI and GLAM estates, and indeed the UK and global economies as a whole, requiring a step-change in infrastructural maturity to provide the foundations for advanced computational research across a very wide range of activity, from History, to Literary Studies, and Archaeology, to keep the UK's enviable world position. The degree of convergence is such that successful large-scale research now often requires translation of computational methods to and from AH across widely divergent domains, including the Cultural Heritage sector but also Heritage Science, Computer Science, Data Science, Software Engineering, Product Development, and the Creative Industries. Recognition of the importance of small-scale digital research, exploiting high quality local datasets or experimenting with delivery of low carbon and minimal computing infrastructure, adds to the richness of the problem domain. The 'applied' research implied by such activity intersects in important ways with major investments by AHRC into the creative industries, requiring RSEs with a wide range of hybrid skills from programming to creative writing.

It is this radical convergence across the arts, humanities, cultural heritage, creative, and technology sectors which offers the most obvious case for investment in AH RSEs. It is broadly accepted that the UK's ability to exploit opportunities created by these convergent domains lies in the development of digital expertise and infrastructural maturity, and AHRC has a core professional community ready-made and actively contributing because of previous investments. The AH RSE community is relatively small (recent surveys by UKRI will hopefully contribute to our understanding), but a major asset that can help deliver key strategic goals for AHRC and enhance the Research Council's ability to contribute to wider UKRI initiatives related to RTP and RSE training, career development, and quality standards.

The complexity of the problem domain, need to align to wider AHRC, UKRI policy initiatives, and relatively recent growth in awareness about the value of AH RSEs, suggests care should be taken not to move too fast. A significant amount can be achieved with relatively modest funding, establishing a stakeholder community and lines of communication,

and putting in place interim governance mechanisms to ensure quality advice can be fed into the wider AHRC and UKRI policy development communities. This would provide a structure to enable additional targeted investment (outlined below) **to enable iDAH in the first instance** and build the foundations for broader and longer-term capacity development aligned to evolving activity across the research, cultural heritage, and creative industry sectors. Appendix 1 (outlined below in Themes) demonstrates the extensive opportunities presented by the existing AH RSE community, but we must recognise that many of them are of a scale of ambition and complexity that careful policy and capability development will be needed to realise them.

The AH RSE community is more than capable of helping AHRC chart a course towards resolving major research challenges and realising a low carbon / environmentally friendly and economically sustainable future, but barriers exist. Many of these are historic and a result of the RSE community (across all Councils) not being as visible as it is now. We accept that they are not necessarily straight-forward to resolve, requiring: reconsideration of funding mechanisms (to better distribute RSE capacity across the UK and incentivise institutions to provide permanent contacts and improve the working conditions of RSEs); changes to research policy (to provide RSEs with the option of acting as PI, ensuring adequate technical review of funding bids, and providing funding for infrastructural development); and development of role descriptions, career paths, and training for staff and students.

These challenges are offset by the active engagement of UKRI in similar matters in cognate roles such as Research Technician, and the substantial analysis and policy development led by SocRSE and SSI, but investment in specifically AH RSE is needed to ensure adequate contributions from AH stakeholders in those initiatives. It should be remembered that AH RSEs often choose a career working with AH and GLAM colleagues ahead of other research domains or careers in the technology sector: they are passionate and informed advocates and allies for the arts and humanities in a time of rapid technological change. AHRC can consider them a community ready-made for the current era, one that has evolved serendipitously through decades of prior investment in Digital Humanities, Digital Cultural Heritage, and Digital Arts. Increasing awareness of their existence across the AHRC and GLAM sector and providing additional investment to build their capability will produce significant short and long term rewards.

Our recommendation is to build on the start documented in this working paper, forming two key committees, reporting to AHRC Head of (Research) Infrastructure, to contribute to AHRC policy development and funding strategy:

1. <u>AHRC RSE Reference Group</u>: A continuation of the current group discussion (see Appendix 4), conducted as a Special Interest Group (SIG) within the UK-EI Digital Humanities Association with modest seed funding provided by AHRC. The purpose would be to provide the AHRC RSE Advisory Group with a forum for discussion, feedback, and where necessary validation.

2. <u>AHRC RSE Advisory Group</u>: A new group of 5-10 people with deep understanding of AH RSE, to act as advisors to AHRC. Their purpose, to be defined in Terms of Reference, would be to provide AHRC with actionable, costed, and validated recommendations for policy, investment, and strategy.

The Advisory Group would as a priority further define, cost, validate, and if necessary extend or reduce the Preliminary Recommendations listed below in consultation with the Reference Group and relevant initiatives such as SocRSE, SSI, UK-EI DH Association, and (where appropriate) international and industry partners. The appendices to this paper would be used as a backlog, informing future analysis.

6. Preliminary Recommendations

The following table lists high-level preliminary recommendations. Refer to Appendix 3: Requirements Definition Matrix for more extensive requirements. Cost ranges have been provided to indicate balance of spend, pending refinement and validation by the Focus Group.

Small: £100k - £1m.
Medium: £1m - £10m.
Large: £10m+

Title	Description	Cost
Governance	1. Establish an AHRC RSE Reference Group, based on the current working group and conducted as a Special Interest Group (SIG) within the UK-EI Digital Humanities Association with modest seed funding provided by AHRC 2. Establish a 5-10 person AHRC RSE Advisory Group, compensated for their time and with a permanent chair, to provide detailed advice to AHRC / iDAH distilled from and validated by the larger Reference Group.	Small (Immediate)

Consultancy	Commission high-level architectural mapping of iDAH scoping documents.	Small (Near Term)
AH RSE Pool	Create a UK AH RSE pool, by funding existing local centres of excellence (freeing their resources for collaboration and capability building) and providing additional funding for 12 RSEs to be hosted in 3-4 institutions across the UK for 10 years. Additional investments should be made to facilitate secondments and the ability to outsource excess demand to a curated community of commercial vendors and non-commercial open source groups.	Medium (Near Term)
Policy	Update AHRC policy to enable AH RSE career development, recruitment & retention, incentivization, research contribution & collaboration, REF alignment, and training.	Small (Near Term)
UK AH RSE Advisory Group	Implement a permanent AH RSE function to review technical proposals, elicit requirements, and define AH / iDAH success metrics, architecture, models (data, processes, workflows), environmental standards, career paths & promotion metrics, training / certification, and technical research standards in alignment to UKRI RSE and infrastructure initiatives. Design policies to support sustainability and archiving of funded research projects, considering models such as ERIC infrastructures CLARIN/DARIAH and with a focus on enabling reusable and environmentally sustainable AH processes, models, data and software. Ensure AHRC - UKRI technical alignment, showcasing challenges and existing case studies to university VCs, BEIS, UKRI heads, peer reviewers. Development of appropriate commercial and non-commercial open source partnerships.	Medium (Near Term)
Training	A range of training initiatives, for AH RSE, AH researchers, and students, with a focus on talent development, capability building, and enhancement of cross-disciplinary and cross-sector collaboration. Ensure alignment with successful skills pilots (2023-). In part, ensure UK-wide funded HPC investments etc. are applicable to and used by AH. Consider augmenting emphasis on teaching interested people to code with teaching a wide range of people how to read code and data, and design and implement algorithms. Consider the skills the AH community has in terms of dealing with ambiguity, missing data, different perspectives and how they can influence data science.	Medium (10 years)
Compute Capability	Enable AH RSE compute capability, leveraging existing UKRI investments and with a view to filling gaps, transferring existing models, and providing documentation and training rather than building new compute infrastructure. Contribute to NetZero DRIC ambition and take account of the media diversity and dimensional complexity of outputs in AH workflows - in experimental form using existing machine provision from other research domains and emphasising modularity, versioning, reuse, sustainability, emerging technology, and enabling challenges. Provide the AH RSE community with the ability to run scalable, reproducible cutting-edge data science and AI workflows.	Small (Near Term)
Challenges	Encourage AH participation in medium and long-term challenge projects related to AI, digital creativity, climate change, quantum computing, and planetary scale data aggregation, management, and analysis. Encourage alignment to and participation with existing AHRC and UKRI-wide challenge funding. Focus on	Large (10 years)

	improving research and infrastructural quality, sustainability, and integration.	

7. Themes

Key themes that emerged across discussion of opportunities, problems, and requirements during the two workshops are detailed in the appendices. They include, but we should assume are not limited to: **Funding (21)**; **Policy (13)**; **Research Challenges (11)**; **Communication (7)**; **Technical (6)**; **Field Building (5)**; **Training (5); Domain Knowledge (5)**; **Research Culture (4); Research Policy (4)**; **Methodological (4); Human Resources (4); Cross-domain Synergies (3)**; **Business (1); Management (1); Competition (1);**.

8. Risks

The following immediate risks have been identified. Ongoing risk management should occur in collaboration between RG and FG.

Title	Description	Priority (H/M/L)	Impact (H/M/L)	Proximity (I/N/F)
Cost estimates	Inadequate requirements definition and validation results in inaccurate costings.	H	H	I
iDAH architecture	Scoping analysis for the iDAH investment does not result in a cohesive design.	H	H	I
RSE identities	People who could benefit from the AH RSE community might not identify as RSEs and so not be heard.	M	M	N
UKRI alignment	AH RSE needs to be aligned to and enabled by wider UK RSE initiatives such as the Society of RSE and SSI.	M	M	I

9. References

Brett, A, Croucher, M, Haines, R, Hettrick, S, Hetherington, J, Stillwell, M & Wyatt, C Research Software Engineers: State of the Nation Report 2017. Engineering and Physical Sciences Research Council (2017). https://doi.org/10.5281/zenodo.495360

Ciula, A., J. Nyhan, and C. Moulin. 'ESF Science Policy Briefing on Research Infrastructures in the Digital Humanities: Landscapes, Ecosystems and Cultures'. *Lexicon Philosophicum* 1 (2013): 277–87.

Ciula, Arianna. 'KDL Checklist for Digital Outputs Assessment', 6 August 2019. https://doi.org/10.5281/zenodo.3361580.

McGillivray, Barbara, Beatrice Alex, Sarah Ames, Guyda Armstrong, David Beavan, Arianna Ciula, Giovanni Colavizza, et al. 'The Challenges and Prospects of the Intersection of Humanities and Data Science: A White Paper from The Alan Turing Institute', 2020. https://doi.org/10.6084/M9.FIGSHARE.12732164.

Professor Mark Parsons, Dr Alastair Basden, Professor Richard Bower, Mr Neil Chue Hong, Professor Davide Constanzo, Mr Shaun de Witt, Professor Luigi Del Debbio, Dr Glen Greed, Dr David Ham, Dr Philip Hasnip, Dr Adrian Hines, Professor Simon Hettrick, Dr Sylvain Laizet, Dr Jeyan Thiyagalingam, Professor Tobias Weinzierl, Professor Garth Wells, Professor Scott Woodley, & Dr Nigel Wood. *ExCALIBUR Research Software Engineer Knowledge Integration Landscape Review.* (2021) https://doi.org/10.5281/zenodo.4986062

Smithies, James, Pascal Flohr, Fadi Balaawi, Sahar Idwan, Carol Palmer, Alessandra Esposito, Shatha Mubaideen, and Shaher Rababeh. '"MaDiH (مديح)": A Transnational Approach to Building Digital Cultural Heritage Capacity'. *ACM Journal on Computing and Cultural Heritage*. Forthcoming. Peer reviewed version (Green OA) (11 January 2022). https://kclpure.kcl.ac.uk/portal/en/publications/madih-(fb4f493f-0a3e-430d-87c2-89aead6df39d).html.

Smithies, James, Patrick Ffrench, and Arianna Ciula. 'Droit de Cité: The Digital Lab as Digital Milieu'. In *Digital Humanities Laboratories: Perspectives on Knowledge, Infrastructure and Culture*, edited by Urszula Pawlicka-Deger and Christopher Thomson. London: Routledge, 2022. Forthcoming.

Smithies, James, Sarah Atkinson, and Elliott Hall. 'Applied Digital Humanities and the Creative Industries in the United Kingdom'. *Digital Humanities Quarterly* 16, no. 2 (2022). http://digitalhumanities.org/dhq/vol/16/2/000628/000628.html.

Smithies, James, and Arianna Ciula. 'Humans in the Loop: Epistemology & Method in King's Digital Lab'. In *Routledge International Handbook of Research Methods in Digital Humanities.*, edited by Kristen Shuster and Stuart Dunn, 155–72. London: Taylor and Francis, 2020.

Smithies, James, Carina Westling, Anna-Maria Sichani, Pam Mellen, and Arianna Ciula. 'Managing 100 Digital Humanities Projects: Digital Scholarship & Archiving in King's Digital Lab'. *Digital Humanities Quarterly* 13, no. 1 (2019).

Smithies, James. 'Research Software (RS) Careers: Generic Learnings from King's Digital Lab', King's College London. Zenodo. 2019. https://doi.org/10.5281/zenodo.2564790.

Software Sustainability Institute (SSI). 'Research Software Engineers', *Software Sustainability Institute*, https://www.software.ac.uk/research-software-engineers. Accessed 07 July 2022.

Appendix 1: Opportunities Definition Matrix

Note: The following matrix should be considered 'draft' and in need of further refinement and validation. It is provided to capture details of workshop conversations, indicate broad areas of interest and priority, and suggest areas for future analysis.

Priority: High / Medium / Low.
Impact: High / Medium / Low.
Proximity: Immediate / Near / Far.
Focus: What opportunities (including blue sky) exist for AH RSEs to unlock the potential for AH-led research to contribute to innovation in non-AH areas (such as but not limited to AI and data visualisation)?
Theme summary: Challenges 11; Research Culture 8; Domain Knowledge 5; Cross-domain Synergies 3; Policy 3; Training 3.

Title	Description	Priority (H/M/L)	Impact (H/M/L)	Proximity (I/N/F)
Cross-domain Synergies (Methods)	Methods developed and validated against AH / GLAM data are likely to transfer well to other domains.	M	M	N
Cross-domain Synergies (Innovation)	Media diversity and dimensional complexity of outputs in AH workflows could prompt innovation if introduced to other domains (cuckoo effect).	M	H	N
Cross-domain Synergies (Data)	Some AH RSEs are integrating and analysing data from widely divergent domains, positioning them to contribute to cross-domain and cross-sector initiatives.	M	M	N
Policy (International)	UK AH RSE policy could be aligned to and inform national/international credit and assessment frameworks (e.g. CASRAI, SSI, RDA).	H	M	I
Policy (AH / iDAH integration)	RSEs could be used to identify threads that connect the many iDAH strands / investments / recommendations / products, and better align academic and technical aspects of AH digital projects generally.	H	H	I
Policy (UKRI alignment)	Alignment with other initiatives in place on RSE related issues could create valuable synergies (data assessment, credit etc) (SSI, DPC, Turing).	M	H	N
Research Culture (AH)	AH RSEs are paid to focus on collaboration, process, method, quality, value rather than only research outputs, making them well suited to advancing UKRI's new research culture goals.	H	H	N
Research Culture (AH)	AH - RSE collaborations force AH researchers to think about collaboration	H	M	N

	in new ways. RSEs can contribute to the definition of refreshed AH research culture, assessment, values.			
Research Culture (AH)	As we enable AH RSE we can also encourage, enable, and train non-technical AH researchers to lead technical projects to ensure research rather than technology leads projects.	L	L	F
Research Culture (Cross-domain)	As we enable AH RSE we can also connect 'traditional' scholars in the humanities with DH practitioners and researchers as well as with scientists to challenge and transform perceived 'dangers' and 'risks' in digital research and provoke a change of paradigm in the ways that lone scholars and traditional researchers work.	M	L	N
Research Culture (Ethics)	AH RSEs have a grounded understanding of historical, ethical and legal questions which underpin data, and the value of context and provenance.	M	H	N
Research Culture (STEM)	Challenges of working with AH data and analysis could contribute to research in a variety of disciplines, and industry and society more broadly.	M	H	N
Research Culture (Data)	AH RSE could help to improve data collection, management, and analysis quality so that datasets are too valuable to be discarded.	H	M	N
Research Culture (Institutional)	Fund and otherwise encourage institutions to hire and use AH RSEs to develop data management and digital research strategies to leverage their specialist knowledge.	H	H	I
Challenges (Infrastructure)	Provide AH RSEs with resources and support (possibly delivered via cross-council and vendor partnerships) related to design architectures, models, workflows, processes for AH digital infrastructure that moves the UK away from point solutions and silos over the medium to long term, enabling advanced methods and translational activity at scale.	H	H	I
Challenges (AI)	Connect common themes or topics across disciplines, for instance, medical / environmental / colonial / source criticism at scale to enable paradigm-breaking AI.	M	M	N

Challenges (Climate Change Methodology)	Complexity of AH data and interpretation, coupled with values-based focus, could contribute to climate change policies and interventions.	M	H	N
Challenges (Climate Change Infrastructure)	Alternatives to new AH large-scale infrastructure should be explored, prioritising low carbon models wherever possible and including minimal computing / reduced energy models / leveraging existing infrastructure etc.	H	M	I
Challenges (Removing Barriers)	Use AH RSE expertise to remove barriers and improve access to AH / GLAM data, models, narratives, policies, workflows and connect global datasets to expand research capability and cultural impact.	H	M	N
Challenge (Knowledge Representation)	Use AH RSE expertise in data modelling, Linked Open Data, and the semantic web to increase access to and connection across UK cultural heritage and research content.	M	H	N
Challenges (Emerging Industries)	Complexity of AH data and interpretation, coupled with focus on creativity and collaboration, could contribute to development of new and emerging industries, products, services.	M	H	N
Challenges (Science Communication)	AH RSE are well placed to communicate the scale, complexity, and value of the human record, the challenges of data science at scale, and the nuances of computational methods.	M	H	N
Challenges (Creative & Cultural Industries)	AH RSEs represent the technical powerhouse of UK digital creative & cultural industries, and are essential to the future of the sector. Experts in transmedia story-telling and multimedia.	M	M	N
Challenges (Quantum Computing)	Opportunity to teach, learn and apply quantum computing algorithms to humanities data (at scale).	M	M	N
Challenges (Data Management)	AH RSEs could contribute to new (non-epistemological) approaches to data management necessary to manage & exploit high dimensionality datasets at scale.	M	H	N
Challenges	AH RSEs could work more closely with	L	L	F

(Cultural Heritage Data)	vendors to change the design of institutional CH systems and metadata which preserve legacy data.			
Challenges (Digitisation)	AH RSEs could work with digitization experts to help with large-scale digitization to enable next-generation methods and use cases (AI, quantum, XR).	L	L	F
Challenges (Born Digital)	AH RSEs are essential to exploiting increasingly large-scale born digital collections, and in some cases designing and maintaining associated infrastructure	M	M	N
Domain Knowledge (Data Science)	Many AH RSEs have deep knowledge of working with historical data that would add value to Data Science (a field realising its dependence on history).	M	M	N
Domain Knowledge (AI)	AH RSE are well placed to transfer AH research perspectives and AH RSE methods to STEM AI research teams, government / commercial teams, and the public.	H	H	I
Domain Knowledge (Commercial)	Methods developed by AH RSEs, informed by deep engagement with socio-cultural and political research, have significant value to commercial sector companies seeking to: profit, manage perception, manage risk, protect users (and exploit users).	M	H	N
Domain Knowledge (Convergence)	AH RSEs work in radically cross-disciplinary contexts (Computer Science - AH) and are expert at converting (technical, social, political) conflicts into convergent solutions.	M	H	N
Domain Knowledge (Reuse)	Data and software created in A&H could be made available for re-use in effective ways, in certified repositories with well-understood licences, at persistent online locations, with high availability, via well-documented APIs, and as open (in FAIR ways) as possible.	H	H	I
Training (Skills Dissemination)	AH RSE are well placed to increase technical awareness and abilities of the AH research community, helping them to read code and data, and transfer AH perspectives to other research domains.	M	H	N
Training (AH)	AH RSE are well placed to augment AH research methods with	M	M	N

		understanding of inference, deduction, modelling, verification and provenance.			
Training (AH RSE)		It is a good time to build awareness of AH RSE into undergraduate and postgraduate education, with a view to defining robust career paths and filling talent gaps.	M	H	I

Appendix 2: Problem Definition Matrix

Note: The following matrix should be considered 'draft' and in need of further refinement and validation. It is provided to capture details of workshop conversations, indicate broad areas of interest and priority, and suggest areas for future analysis.

Priority: High / Medium / Low.
Impact: High / Medium / Low.
Proximity: Immediate / Near / Far.
Focus: What barriers and challenges make it difficult for AH RSEs to unlock the potential for AH-led research to contribute to innovation in non-AH areas (such as but not limited to AI and data visualisation)?
Theme summary: Technical 6; Methodological 4; HR 4; Management 1; Competition 1; Research Policy 4; Field Building 7; Funding 6; Communication 5; Training 2; Business 1.

Ref.	Title / Theme	Description	Priority (H/M/L)	Impact (H/M/L)	Proximity (I/N/F)
P001	Competition	GLAMs are at a disadvantage to universities and both are at a disadvantage to the commercial sector.	H	H	I
P002	RTP identities	RSE and Research Technical Professional (RTP) identities overlap, complicating policy development and field building.	M	L	N
	Management (Contract Status)	Short & fixed term contracts prevent long-term engagement & capacity building, infrastructure maintenance / development, and discourage the choice of an RSE career path.	H	H	I
	Research Policy (Collaborative credit)	Collaborative research is encouraged but single author outputs are sometimes needed for career advancement. REF credit needs updating to include GitHub etc. Non-technical researchers need education & policy guidance.	H	M	N
	Research Policy (PI status)	There are barriers to non-academic staff acting as PI on grants. Methods of overhead	H	M	I

		recovery of RSEs are variable depending on institution / contract type.			
	HR (Recruitment)	It is difficult recruiting for AH RSE roles, even with funding. Institutions are not offering and funders are not enabling long-term, stable roles, career paths, promotion processes, salary benchmarking	H	H	I
	HR (Career pathways)	Undergraduate, postgraduate, and career pathways & incentives for RSEs (analysts, designers, engineers etc) do not exist. Potential academic pathways & incentives for RSEs are not optimised or used. Commercial pathways are more attractive.	H	M	I
	HR (Community)	Long-standing unrecognised AH RSE activity are not integrated into / contributing to UK RSE initiatives and communities. RSEs working alone, in pools, or in groups have different experiences and needs.	L	L	N
	Funding (Technical)	There is a lack of funding for devoted AH RSE work, unlike in EPSRC and STFC. Funding prioritises non-technical outcomes and views RSE work as instrumental / secondary.	H	M	I
	Funding (Institutional)	RSE talent is spread across institutions but funding incentivizes institutions to resist pan-institutional RSE collaboration and act competitively.	M	M	N
	Funding (Focus)	Projects reflect funder language, which in AH does not include or enable RSE.	M	M	N
	Funding (Innovation)	Funding for early stage innovation and experimentation is lacking, and dependent on institutional resources / culture. This is compounded by lack of funder clarity about RSE roles in projects (service provider vs researcher).	H	M	I
	Funding (Architecture & Integration)	There are a lack of funding schemes that look beyond point solutions to design, develop and	H	H	I

		sustain RSE staff / systems / software / models / methods / processes.			
	Funding (Optimisation & Sustainability)	There are a lack of funding schemes, administrative support, and funding pathways that support optimisation / scaling up / care / maintenance of what has already been built. Funders leave AH RSEs with a 'hospital pass' of their abandoned investments.	H	M	I
	Communication (Role Purpose)	There is a of understanding about what the RSE role is across, and is capable of doing, outside the RSE community.	H	M	I
	Communication (Transferability)	AH is not seen as a source of transferable research and skills by policy and sector stakeholders.	M	M	N
	Communication (Perception)	The results of AH innovation & data science could be better communicated to non-AH audiences.	M	L	N
	Communication (Technical)	Interdisciplinary and cross-sector communication is hampered by lack of mutual understanding, and complicated by the technical nature of RSE roles.	L	M	N
	Communication (Interdisciplinary)	Disciplinary patch protection leads to disciplinary cultural & linguistic barriers within RSE just as with research disciplines.	M	M	N
	Communication (Data & use case complexity)	Research domains that might benefit from exposure to AH data complexity, standards, and use case challenges have not been incentivized to collaborate.	M	M	N
	Field Building (Intellectual)	The intellectual value of RSEs is not recognised by researchers or policy makers. This is worsened by a lack of clear intellectual collaboration guidelines to maximise potential. RSEs are usually forced into service roles or (less commonly) forced into intellectual roles. Communication of possible approaches, and flexible policies are needed.	M	H	M
	Field Building	The system undervalues practical	M	L	N

	(Applied Skills)	/ applied skills. There is a lack of recognition of the value of even basic coding skills.			
	Field Building (Hiring)	RSE leaders are reluctant to hire people from Humanities backgrounds.	M	M	N
	Field Building (Attribution)	There is a reluctance to add RSEs as authors. REF credit needs updating, adding flexibility to ensure reward can match motivation. There is a lack of understanding about AH-specific RSE assessment metrics / standards.	M	M	N
	Field Building (AH Technical Research)	There is a long-standing stigma against AH technical research in AH, often based in ignorance. This results in a lack of understanding of the variation of technical RSE skills and needs across AH disciplines.	M	M	N
	Field Building (STEM)	Greater communication of the challenges and results of ML/AI research across research domains is needed, to spur innovation and avoid naive adoption of techniques.	M	H	N
	Field Building (Failure)	The potential inherent in failure is not embraced, leading to risk aversion and lack of transparency.	M	M	N
	Field Building (Barriers to access)	Most researchers struggle to find RSE support, and it tends to be time limited.	H	M	N
	Field Building (Translational Research)	Translation from Comp Sci research to AH / GLAM methods requires investment and improved communication.	M	M	N
	Policy (Assessment)	There are a lack of standards for technical assessment of funded projects and research outputs.	H	H	I
	Methodology (Ontologies & Standards)	The complexity of AH domains mean there are no cross-domain ontologies or standards, either within AH or outside. Extends to workflows, metadata, interoperability, etc.	M	M	N
	Policy (Systems vs Data)	There is a lack of awareness and policy / funding support for the difference between data and	M	M	N

		software / systems reduces innovation, quality, reproducibility, and sustainability.			
	Methodological (Data management)	Static rather than dynamic approaches to data are incentivised, compounded by a lack of relational theory.	M	M	N
	Technical (IAAS)	Managed and agnostic service models (IAAS) are needed for infrastructure provision, including access to HPC / scalable compute.	M	M	N
	Technical (Architecture)	Integrated, agnostic, and modular architecture needs to be defined and maintained, to facilitate stand-alone innovation and experimentation, UKRI alignment, and benefits of scale. Enterprise as well as solution / system architecture is needed.	H	H	I
	Technical (AH complexity)	AH requires that computers are able to deal with gaps and biases in data and encoding, uncertainty, ambiguity, complexity, and competing interpretations, pushing them to the limits of their capabilities.	M	M	N
	Technical (Sustainability)	Sustainability funding is limited, resulting in security risks and risk-averse or limited solutions.	H	H	I
	Technical (Documentation)	A repository of technical information, best practice, methods etc is needed. E.g; The Turing Way.	M	M	N
	Methodological (Domain Knowledge)	Assumption that RSEs can always cross research domains ('coding is cross-disciplinary') ignores the importance of (technical and non-technical) domain knowledge. Need for AH-specific RSE pool.	M	M	N
	(Methodology) Disciplinary integration	Technical knowledge is not properly integrated into AH research, leading to policy and quality issues.	M	L	I
	Methodology (Design)	Building outputs without a user base, rather than via practitioner-research methods; building knowledge not reference systems.	L	L	F

	Business (Cross-sector interests)	Commercial and AH GLAM sectors have different motivations, leading to siloed systems.	M	M	N
	Training (Critical Thinking)	Tools are taught at the expense of critical / epistemological understanding and core skills such as modelling, humanities research methods.	M	H	N
	Training (Tools)	AH training is needed, to introduce tools, share advanced DH methods, upskill students and self-taught researchers, and ensure colleagues are able to work effectively with RSEs	H	H	N

Appendix 3: Requirements Definition Matrix

Note: The following matrix should be considered 'draft' and in need of further refinement and validation. It is provided to capture details of workshop conversations, indicate broad areas of interest and priority, and suggest areas for future analysis.

Priority: High / Medium / Low.
Impact: High / Medium / Low.
Proximity: Immediate / Near / Far.
Focus: What are the SMART requirements (technical, human, policy, funding) for AH RSEs to unlock the potential for AH-led research to contribute to innovation in non-AH areas (such as but not limited to AI and data visualisation)?
Themes: Funding 15; Policy 10.

Description	Priority (H/M/L)	Impact (H/M/L)
Implement a permanent AH RSE function to review technical proposals, elicit requirements, and define AH / iDAH success metrics, architecture, models (data, processes, workflows), environmental standards, career paths & promotion metrics, training / certification, and technical research standards in alignment to UKRI RSE and infrastructure initiatives, and considering models such as CLARIN/DARIAH and with a focus on enabling reusable and sustainable AH data and software. Also responsible for AHRC - UKRI alignment and advocacy, showcasing grand challenges and existing case studies to university VCs, BEIS, UKRI heads, peers reviewers. Aligned to and enabling of appropriate commercial and open-source non-commercial partners.	H	H
Create a UK AH RSE pool, by funding existing local centres of excellence (freeing their resources for collaboration and capability building) and adding	H	H

possibility of secondments and additional central capacity alongside a curated community of commercial vendors.		
Fund the scoping of an AH RSE compute capability based on existing cross-Council investments, taking account of the media diversity and dimensional complexity of outputs in AH workflows - in experimental form using existing machine provision from other research domains and emphasising modularity, versioning, reuse, sustainability, emerging technology, and enabling grand challenges.	M	H
Provide seed funding for concept development and consortium building for 4 challenges related to infrastructure, climate change, quantum computing, AI, or digital creativity.	M	M
Fund medium and long-term challenge projects, also tasked with improving quality, sustainability, integration etc.	M	M
Provide major funding tranches to digitise 'everything' to FAIR standards, to enable next-generation methods and use cases (AI, quantum, XR).	L	M
Fund cross-Council RSE champions for 5 years, to transfer insights, attitudes, and methods across research domains and sectors.	M	L
Fund cross-Council RSE champions for 5 years, to enable technology transfer from AH RSE to other research domains, government, commercial sector.	M	L
Equivalent funding programs to https://www.ukri.org/opportunity/software-for-research-communities/ & https://www.ukri.org/opportunity/support-the-development-of-research-software-engineering/ focused on AH and its applications to non-AH remits.	H	H
Fund cross-disciplinary training in digital concept development leveraging different traditions of innovation, and high-risk projects that might fail..	M	H
Invest in a long term training programme covering aspects of AH RSE and DH, with a train the trainer model.	H	H
Build a coordinated training capability within AH to upskill the AH community in AI and RSE skills.	M	L
Offer training in cross-domain communication, for all researchers and disciplines.	M	M

Fund development of UK-wide AH RSE undergraduate and postgraduate career pathway definition.	H	H
Funding for regional and national network activities and hackathons, encouraging interdisciplinary collaborations (for example, bringing together AI expertise and AH researchers).	M	L
Use RSEs in peer review processes to ensure technical quality, paying their institutions for their time.Consider reinstating AHRC technical review process.	H	H
Work with institutions to develop and deploy policies to enable AH RSE career training, pathways, incentives. Reach into schools, as with https://www.stem.org.uk/stem-ambassadors.	H	H
Incentivise or instruct institutions and funders to position RSEs as PIs and Co-Is and /or ensure well defined (named) project roles for them, and audit them to ensure buy-out is provided with successful funding.	H	M
Involve AH RSEs in funding policy development processes, and ensure RSE outcomes are incentivised in funding opportunities (including experimentation, modelling, architecture, sustainability).	H	H
Ensure environmental considerations for conducting digital research are taught as a requirement for all AH and RDM practitioners, tool builders, and service providers, hardware vendors.	H	M
Work with UKRI, RSESoc, SSI to develop a UK RSE policy aligned to national/international credit and assessment frameworks (e.g. CASRAI, SSI, RDA).	H	M
Incentivise, value, and measure quality of RSE contributions to funded research outputs.	H	H
Avoid zero-sum competition between AH RSE and academics by maximising potential to gain AH RSE funding from non-AHRC sources.	M	M
Increase AH RSE involvement in the Society of RSE and ensure representation at SocRSE trustee board: 1 by 2025; 2 by 2027.	H	H
Continue to include software creation as an impact metric for REF, feeding into DataCite (DataCite software DOI). Encourage software citation and incentivise reproducibility.	H	M

Appendix 4: Members

Role	Member	Institution
Chair	James Smithies	King's College London
AHRC - UKRI representatives	Tao-Tao Chang	AHRC
	Anna-Maria Sichani	AHRC / University of London
	James Fenner	AHRC
	Melodee Wood	AHRC / Loughborough University
Members	Neil Jefferies	University of Oxford
	Dave de Roure	University of Oxford
	Lorna Hughes	University of Glasgow
	Charlotte Tupman	University of Exeter
	Victoria Moody	JISC
	Martin Steer	SAS
	Jane Winters	SAS
	Ruth Ahnert	QMUL
	Sharon Webb	University of Sussex
	Pip Willcox	The National Archives
	John Moore	The National Archives
	Mia Ridge	British Library
	Filipe Bento	British Library
	Dominic Oldman	British Museum
	David Beavan	Alan Turing Institute
	Pieter Francois	Alan Turing Institute
	Timothy Hobson	Alan Turing Institute
	Arianna Ciula	King's College London
	Barbara McGillivray	King's College London
	Patricia Murrieta-Flores	Lancaster University
	Tiago Sousa Garcia	Newcastle University
	Mark Turner	Newcastle University
	Melissa Terras	University of Edinburgh
	Lisa Otty	University of Edinburgh

	Marion Weinzierl	Durham University / N8 CIR
	Julianne Nyhan	University College London
	Raquel Alegre	University College London
	Anne Alexander	University of Cambridge
	Mary Chester-Kadwell	University of Cambridge
	James Hetherington	University College London
	James Baker	University of Southampton
	Steve Crouch	University of Southampton
	Tim Evans	Archaeology Data Service, University of York
	Martin Wynne	Oxford Text Archive