

REGIONAL HETEROGENEITY IN UKRI FUNDED PROJECTS

Exploring regional innovation clusters
using text-mining approaches

IRC Report No: 091

REPORT PREPARED BY

Qijun Zhou
University of Glasgow

Nuran Acur
University of Glasgow

Authors

The core members of the research team for this project were as follows:

- » Dr Qijun Zhou - University of Glasgow
- » Dr Nuran Acur - University of Glasgow

This document relates to IRC Project FFGtR007: Regional heterogeneity and influence of UKRI-funded projects: an investigation using topic modelling approaches

Acknowledgements

This work was supported by Economic and Social Research Council (ESRC) grant ES/X010759/1 to the Innovation and Research Caucus (IRC).

We are very grateful to the project sponsors at UK Research & Innovation (UKRI) for their input into this research. The interpretations and opinions within this report are those of the authors and may not reflect the policy positions of UKRI or its constituent councils. We would also like to acknowledge and appreciate the efforts of the IRC Project Administration Team for their support in preparing this report.

About the Innovation and Research Caucus

The Innovation and Research Caucus supports the use of robust evidence and insights in UKRI's strategies and investments, as well as undertaking a co-produced programme of research. Our members are leading academics from across the social sciences, other disciplines and sectors, who are engaged in different aspects of innovation and research systems. We connect academic experts, UKRI, IUK and the (ESRC), by providing research insights to inform policy and practice. Professor Tim Vorley and Professor Stephen Roper are Co-Directors. The IRC is funded by UKRI via the ESRC and IUK, grant number ES/X010759/1. The support of the funders is acknowledged. The views expressed in this piece are those of the authors and do not necessarily represent those of the funders.

Contact

You are also welcome to email us if you have any questions about this report or the work of the IRC generally: info@ircaucus.ac.uk

Cite as: Zhou, Q. & Acur, N. 2026. *Regional Heterogeneity in UKRI Funded Projects. Exploring Regional Innovation Clusters Using Text-Mining Approaches* Oxford, UK: Innovation and Research Caucus

CONTENTS

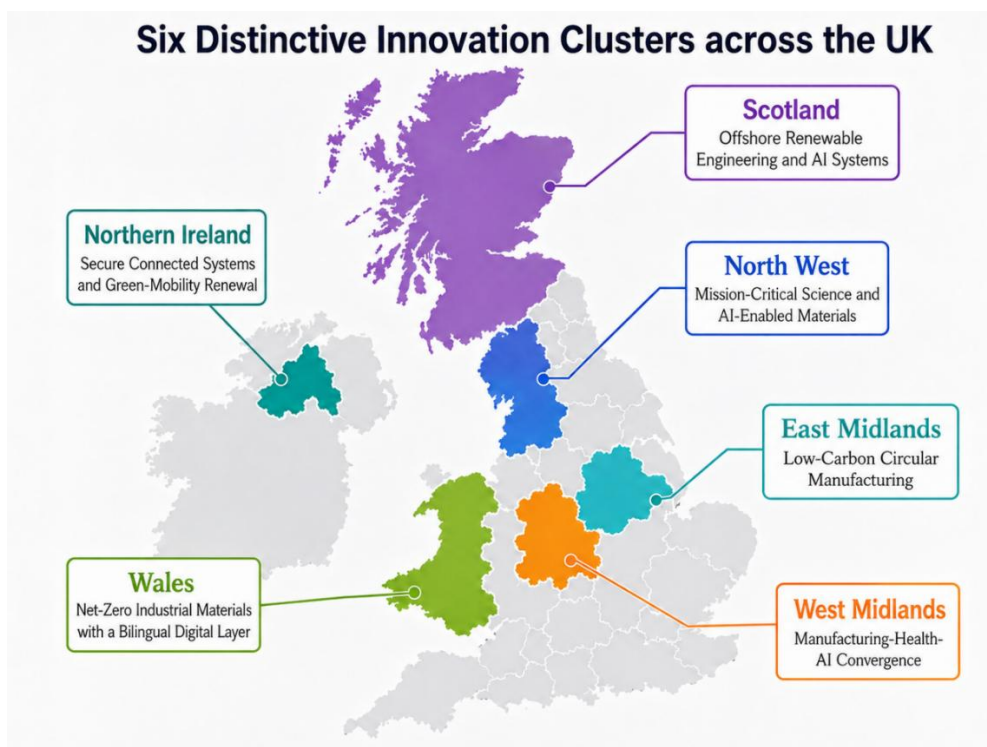
EXECUTIVE SUMMARY	4
1. INTRODUCTION	6
1.1 The evolution of the UK place-based innovation regime	6
1.1.1 Stage 1: national excellence and knowledge exchange (2006–2016).....	6
1.1.2 Stage 2: Industrial Strategy and explicit place-based additionality (2017–2020)	6
1.1.3 Stage 3: Place becomes a formal strategic objective (2021–2022).....	7
1.1.4 Stage 4: Local innovation cluster pilots (2022–2024)	7
1.1.5 Stage 5: Locally led, portfolio-scale cluster investment (2025–2031).....	7
1.2 Research Aim and Questions	8
2. METHODS	9
2.1 Analytical approach	9
2.2 Text corpus and data structure	9
2.3 Text preprocessing	9
2.4 BERTopic modelling procedure	10
2.5 Topic interpretation and validation	11
2.6 Analysis of topic distributions across institutions and time	11
3. RESULTS.....	11
3.1 Top topics appeared in UKRI-funded projects	11
3.1.1 Topic 0: Society, policy, culture and gender.....	11
3.1.2 Topic 1: Climate, plants, crops and species.....	12
3.1.3 Topic 2: Cells, cancer, proteins and infection.....	12
3.1.4 Topic 3: Quantum technologies, devices, chemical reactions and batteries	12
3.1.5 Topic 5: Brain, neurons, memory and neural systems.....	13
3.1.6 Topic 4: Galaxies, stars, the universe and physics	13
3.1.7 Topic 9: Security, business, services and users	13
3.1.8 Topic 11: Care, patients and healthcare	13
3.1.9 Topic 6: Equations, mathematics and geometry	14
3.1.10 Topic 7: Music, games and sound.....	14
3.2 Hierarchical clustering	14
3.3 Similarity matrix	16
4. THEMATIC DIFFERENCES ACROSS UK REGIONS.....	17
4.1 East Midlands	17
4.2 East of England	18
4.3 London.....	19
4.4 North East.....	20
4.5 North West.....	21
4.6 Northern Ireland.....	22
4.7 Scotland.....	23
4.8 South East.....	24
4.9 South West.....	25
4.10 Wales.....	26

4.11 West Midlands	28
4.12 Yorkshire and the Humber	29
5. EVOLUTION OF THEMATIC HETEROGENEITY	31
6. DISCUSSION: IDENTIFYING SIX INNOVATION CLUSTERS ACROSS THE UK.....	33
6.1 East Midlands - Low-Carbon Circular Manufacturing Regional innovation cluster.....	33
6.2 North West - Mission-Critical Science and AI-Enabled Materials Regional innovation cluster	34
6.3 Scotland - Offshore Renewable Engineering and AI Systems Regional innovation cluster	35
6.4 Wales - Net-Zero Industrial Materials with a Bilingual Digital Layer	36
6.5 West Midlands - Manufacturing-Health-AI Convergence Regional innovation cluster	37
6.6 Northern Ireland - Secure Connected Systems and Green-Mobility Renewal	38
7. CONCLUDING REMARKS	38
7.1 Limitations and future research	39
7.2 Policy implications	39
REFERENCES	41
APPENDIX	48
Appendix A. Full topic descriptions	48
Appendix B. Top 10 Topics per Region	54
Appendix C. Thematic Evolution per Region	60

EXECUTIVE SUMMARY

For nearly two decades, UK innovation policy has progressively narrowed its focus from individual institutions to place: the Strength in Places Fund, the Levelling Up White Paper, UKRI's Strategy 2022–27 and, most recently, the £500 million Local Innovation Partnerships Fund have all bet that funding can be steered towards distinct regional research strengths. This report asks whether that bet is paying off. Using BERTopic, an AI-based topic-modelling method, we analysed the abstracts of every UKRI-funded project in the GtR+ database, identifying 69 underlying research themes and tracking how they are distributed across the 12 UK regions and across time (2006–2024).

The headline finding represent a paradox. At the national level, funding is concentrated and stable within four broad themes: 1) society, policy, culture and gender; 2) climate, plants, crops and species; 3) cells, cancer, proteins and infection; and 4) quantum technologies, devices and batteries. This national priority has also been reflected on the regional level. Across the 12 regions examined in this study, the top 4 most funded themes are in some combination of the four national priorities. In addition, this core structured was relatively stable since 2006. — dominate the portfolio in almost every region, and this core structure has barely moved since 2006 (the top ten topics' combined share fell only from 64.4% to 60.4%). The genuine regional distinctiveness appeared in what we called 'secondary specialisation', which can be observed in funded topics outside the top 4. These secondary specialisations are also — particularly reflected in university cities and industrial ecosystems.



Reading these secondary specialisations together, the report identifies six distinctive, if still-maturing, innovation clusters:

- » Low-carbon circular manufacturing in the East Midlands (Nottingham/Loughborough hydrogen, plastics and metals);
- » Mission-critical science and AI-enabled materials in the North West (Manchester/Liverpool nuclear and astronomy);
- » Offshore renewable engineering and AI systems in Scotland (Edinburgh/Glasgow wind and language technology);
- » Net-zero industrial materials with a bilingual digital layer in Wales;
- » Manufacturing-health-AI convergence in the West Midlands (Coventry/Warwick/Birmingham); and
- » Secure connected systems with green-mobility renewal in Northern Ireland (Belfast).

Rather than marking new regional trajectories, each cluster represents an existing industrial or scientific strength being recombined with new technologies — and most remain at an early or intermediate stage of maturity.

These findings carry direct implications for the design of the next generation of place-based funding. First, Local Innovation Partnerships Fund awards should be targeted at these secondary specialisations rather than the shared national themes that already attract the bulk of mainstream UKRI funding. Second, regions should map their specialisations explicitly against the Modern Industrial Strategy's eight sector plans to strengthen the case for co-investment. Third, funding models should be differentiated by cluster maturity: established clusters such as Scotland's offshore-wind ecosystem are ready for scale-up finance and stronger private-match conditions, while emerging clusters such as Northern Ireland's secure-systems base still need earlier-stage support to build institutional connections and credible use cases. The central policy challenge, in short, is no longer identifying where research strength exists — it is building the coordination, infrastructure and adoption mechanisms needed to convert that strength into durable, regionally anchored economic benefit.

1. Introduction

Addressing regional inequalities and driving place-based growth and productivity has been a strategic priority for UKRI policy over the past two decades. From the early emphasis on national research excellence and institutional knowledge exchange, through the introduction of explicit place-based additionality tests under the Strength in Places Fund, to the more recent consolidation of place as a formal strategic objective within UKRI's corporate strategy, policy has progressively shifted its unit of analysis from the individual institution towards the regional or functional cluster. This strategic focus creates a pressing need to understand whether UKRI's funding patterns effectively steer different regions toward distinct technological or thematic specialisations. However, the place-based innovation regime governing UKRI funding has not emerged from a single policy decision. Thus, we first present a gradual evolution of the policy focus spanning almost two decades as the background of our study.

1.1 The evolution of the UK place-based innovation regime

1.1.1 Stage 1: national excellence and knowledge exchange (2006–2016)

Stage 1 describes the policy inheritance from which UKRI later emerged. Policy prioritised the quality and connectivity of the national research base and the translation of knowledge into economic and public value. Innovation Nation articulated a broad national innovation agenda; the 2014 science and innovation plan reinforced research excellence, infrastructure and collaboration; the Nurse Review focused on the coherence of the research-council system; and the Higher Education Innovation Fund supported recurrent institutional knowledge exchange (Department for Business, Enterprise and Regulatory Reform, 2008; HM Treasury & Department for Business, Innovation and Skills, 2014; Department for Business, Innovation and Skills, 2015; Research England, 2026b). At this initial stage, locally governed regional innovation clusters were not yet the principal unit of funding design.

1.1.2 Stage 2: Industrial Strategy and explicit place-based additionality (2017–2020)

The 2017 Industrial Strategy incorporated local economic strengths and Local Industrial Strategies into the national growth architecture. Following UKRI's establishment in 2018, the Strength in Places Fund (SIPF) introduced an explicit place-based additionality test: proposals were required to combine research and innovation excellence with credible business participation and additional economic benefit within a self-defined functional geography. Its funded portfolio and evaluation framework made collaboration, counterfactual additionality and place-level outcomes explicit design concerns (Department for Business, Energy and Industrial Strategy, 2017; UKRI, 2023a; UKRI, 2021; Research England, 2021). The 2020 UK Research and Development Roadmap subsequently committed government to a UK R&D Place Strategy. It linked research investment to levelling up (Department for Business, Energy and Industrial Strategy, 2020). The policy unit had therefore shifted from institution-level translation towards a functional cluster proposition.

1.1.3 Stage 3: Place becomes a formal strategic objective (2021–2022)

The UK Innovation Strategy consolidated clusters and regional strengths as components of national innovation policy. At the same time, the Levelling Up White Paper supplied a wider cross-government rationale for spatial distribution and local growth (Department for Business, Energy and Industrial Strategy, 2021; Department for Levelling Up, Housing and Communities, 2022). Most importantly, the UKRI Strategy 2022-27 made place one of six strategic objectives and committed the organisation to strengthening clusters and partnerships throughout the UK. The corporate plan and its subsequent update translated that objective into organisational commitments (UKRI, 2022b; UKRI, 2022a; UKRI, 2023b). This stage institutionalised a strategic objective before locally co-designed delivery had been demonstrated at scale.

The strategic turn was also distributed across UKRI's constituent bodies. Delivery plans specified place-relevant responsibilities for Innovate UK, Research England and the councils supporting engineering, social science and environmental research (Innovate UK, 2022; Research England, 2022a; EPSRC, 2022; ESRC, 2022b; NERC, 2022), as well as for the arts and humanities, biosciences, medical research and large scientific facilities (AHRC, 2022; BBSRC, 2022; MRC, 2022; STFC, 2022). This broad institutional alignment matters because a place-based portfolio depends on coordination across disciplinary, translational and business-facing instruments rather than on a single regional fund.

1.1.4 Stage 4: Local innovation cluster pilots (2022–2024)

UKRI and the government operationalised the place objective through a portfolio of experimental instruments. Innovation Accelerators enabled locally shaped portfolios in Greater Manchester, the West Midlands and Glasgow City Region. Place-Based Impact Acceleration Accounts supported research organisations to construct local impact partnerships; the Regional Innovation Fund expanded devolved knowledge-exchange capacity; and Local Policy Innovation Partnerships strengthened local evidence and policy capability (UKRI, 2026d; UKRI, 2024a; Research England, 2023; ESRC, 2022a). The Connecting Capability Fund provided a complementary mechanism for inter-institutional collaboration (Research England, 2026a). At the same time, the UK Science and Technology Framework and Investment Zones connected local experimentation to wider technology and growth priorities (Department for Science, Innovation and Technology, 2023; Department for Levelling Up, Housing and Communities & HM Treasury, 2023). The substantive change was therefore not merely an increase in regional expenditure, but greater local participation in diagnosis, portfolio selection, coordination and learning.

1.1.5 Stage 5: Locally led, portfolio-scale cluster investment (2025–2031)

The 2025 Modern Industrial Strategy connected regional clusters to eight national growth-driving sectors, detailed through sector plans and an updated Science and Technology Framework (Department for Business and Trade, 2025b; Department for Business and Trade,

2025a; Department for Science, Innovation and Technology, 2025b). The English devolution settlement and the Industrial Strategy Zones Action Plan further strengthened the role of territorial institutions in growth and investment governance (Ministry of Housing, Communities and Local Government, 2024; Ministry of Housing, Communities and Local Government, 2025). Within this architecture, the Local Innovation Partnerships Fund (LIPF) formalised locally led triple-helix partnerships and portfolio construction for high-potential existing and emerging clusters. Its completed strand, selected areas and initial projects, mark the transition from programme design to early implementation (UKRI, 2026e; UKRI, 2025b; UKRI, 2026g; UKRI, 2026b). The 2026-31 UKRI Strategy, corporate-plan update, Delivery Plan and objectives integrate this place mechanism with priority programmes, company growth and devolution (UKRI, 2026j; UKRI, 2025c; UKRI, 2026h; UKRI, 2026f). Higher Education Innovation Funding priorities extend the same long-term orientation within the knowledge-exchange system (Research England, 2025b). Stage 5 thus denotes portfolio-scale policy design; it does not yet imply demonstrated regional transformation.

1.2 Research Aim and Questions

Despite this sustained policy attention, systematic evidence remains limited on whether UKRI's funding produces the regionally differentiated research portfolios that place-based policy aspires to create. Understanding whether distinct technological or thematic specialisations are visible within the pattern of funded projects is therefore essential both for evaluating the effectiveness of place-based instruments to date and for informing the design of the next generation of cluster investments. This report addresses this gap by applying a text-based, data-driven approach to the GtR+ database containing UKRI-funded project abstracts.

Accordingly, this report aims to explore whether, under the evolving place-based innovation policy regime, early signs of geographically distinct innovation clusters can be evidenced through the pattern of funded research projects. Using topic modelling, this report identifies the principal research themes represented in the funding portfolio. It examines how these themes are distributed across UK regions and how that distribution has evolved. The following two research questions are proposed:

- RQ1: Do UKRI-funded projects exhibit significant thematic differences across UK regions?
- RQ2: How does regional thematic heterogeneity evolve?

2. Methods

2.1 Analytical approach

We use topic modelling to identify the principal research areas represented in UKRI-funded projects. Topic modelling is an inductive text analysis method that identifies recurring patterns in a corpus and reveals the thematic structure underlying a large collection of documents (Blei et al., 2003; Grimmer & Stewart, 2013). It is increasingly used in management research to identify phenomenon-based constructs and grounded conceptual relationships from textual data (Hannigan et al., 2019). The underlying premise is that documents addressing related subjects tend to exhibit systematic similarities in their linguistic content. By identifying these recurring patterns across a corpus, topic modelling enables researchers to analyse thematic structures that may be difficult to detect consistently through manual coding alone (Grimmer & Stewart, 2013).

We apply BERTopic, an unsupervised topic modelling method that combines contextualised document embeddings, dimensionality reduction, density-based clustering, and class-based term weighting (Grootendorst, 2022). BERTopic is suitable for the present study because funded-project abstracts may describe substantively related research areas using different words and expressions. Traditional probabilistic topic models, such as latent Dirichlet allocation, primarily infer topics from patterns of word occurrence across documents (Blei et al., 2003). BERTopic complements this word-based information with transformer-based document embeddings that capture contextualised semantic relationships between texts (Grootendorst, 2022; Reimers & Gurevych, 2019). This approach allows the model to identify projects with related substantive content even when their abstracts do not use identical vocabularies.

2.2 Text corpus and data structure

Our corpus comprised the abstracts of all funded research projects included in our sample. We treated each project as a separate textual unit. For each project, we retained five fields: the recipient university name, the university's region, the project title, the project abstract, and the year the project received funding. We used the project abstract as the primary textual input for topic modelling. We retained the remaining fields as metadata and used them after topic identification to examine how topic distributions varied across universities, regions, and years.

2.3 Text preprocessing

Before estimating the model, we standardise the textual data to reduce formatting inconsistencies. We convert all text to lowercase, remove punctuation marks, and normalise whitespace. At this stage, we retain the words contained in each abstract because transformer-based embedding models rely on the contextual relationships among words to generate semantically meaningful document representations.

We apply more targeted vocabulary filtering during the topic representation stage. After the document embeddings are generated and the documents are assigned to clusters, BERTopic uses a class-based term frequency–inverse document frequency procedure, or c-TF-IDF, to identify the terms that distinguish each topic from the other topics in the corpus. At this stage, we remove standard English stop words and apply a customised stop-word list to exclude recurring terms that provide limited information for interpreting the substantive focus of the topics. Examples include general research-related terms such as “project,” “research,” “study,” “data,” “method,” “approach,” “analysis,” “result,” “programme,” and “proposal.” This separation allows the embedding model to retain the contextual information required to identify semantically related projects while reducing noise in the terms used to interpret and label the resulting topics.

2.4 BERTopic modelling procedure

The BERTopic analysis proceeds in four stages. First, we convert each project abstract into a contextualised semantic embedding. We use the all-mpnet-base-v2 model provided by Sentence Transformers. This model maps each sentence or paragraph into a 768-dimensional dense vector representation and is suitable for clustering and semantic similarity analysis (Reimers & Gurevych, 2019). The resulting embeddings capture semantic relationships between documents and place abstracts with related meanings close together in the vector space, even when they use different words (Grootendorst, 2022).

Second, we apply Uniform Manifold Approximation and Projection, or UMAP, to reduce the dimensionality of the document embeddings. UMAP is a manifold-learning technique for nonlinear dimensionality reduction. It reduces the dimensionality of the embedding space while preserving the data's underlying structure (McInnes et al., 2018). This step produces a lower-dimensional representation that facilitates the subsequent clustering analysis.

Third, we use Hierarchical Density-Based Spatial Clustering of Applications with Noise (HDBSCAN), to group semantically related project abstracts into topic clusters. HDBSCAN identifies clusters by examining the density structure of the reduced embedding space and constructing a hierarchy of density-based clusters (Campello et al., 2013). Unlike clustering procedures that require specifying the number of clusters in advance, HDBSCAN derives clusters from the data distribution. It also allows documents that do not fit clearly within a cluster to remain unassigned as outliers (Campello et al., 2015).

Fourth, we use class-based term frequency–inverse document frequency, or c-TF-IDF, to identify the terms that are most representative of each topic. This procedure combines the documents assigned to the same cluster into a single topic-level document. It assigns greater weight to terms that occur frequently within one topic but less frequently across other topics (Grootendorst, 2022). The resulting terms provide a concise representation of each topic's substantive content and support the subsequent interpretation and labelling of the topic clusters.

2.5 Topic interpretation and validation

The topic labels are not assigned automatically. We interpret each topic by reviewing its highest-scoring c-TF-IDF terms, its most representative project abstracts, and the titles of the projects assigned to the cluster. Two researchers independently review the model output and assign a descriptive label to each topic. Where the initial labels differed, the researchers discuss the underlying texts and agree on a final label. This procedure ensures that the topic labels reflect the substantive meaning of the underlying projects rather than relying solely on individual keywords.

2.6 Analysis of topic distributions across institutions and time

After identifying and validating the topics, we assign each project abstract to its corresponding topic cluster. We then calculate the prevalence of each topic across universities, regions, and years. To account for differences in the total number of funded projects across institutions and periods, we measure topic prevalence as the proportion of projects associated with a given topic within each university, region, or year. This analysis allows us to examine which research areas received greater attention in different institutional and geographical contexts. It also enables us to trace changes in topic prevalence over time and to identify whether particular research areas expanded, declined, or remained stable during the study period.

3. Results

Using BERTopic, we identified 69 topics from the corpus, suggesting that the research field exhibits substantial thematic diversity and a broad knowledge structure. Here, we discuss the top 10 topics with the most appearances. The full list of topics is provided in Appendix A.

3.1 Top topics appeared in UKRI-funded projects

The topic-modelling results reveal a highly diverse UKRI-funded research portfolio spanning the social sciences and humanities, environmental and agricultural research, biomedical science, physical science, mathematics, digital technologies, healthcare and the creative industries. The following descriptions interpret the principal substantive directions represented by the most prevalent topics.

3.1.1 Topic 0: Society, policy, culture and gender

Topic 0 is the largest, containing 18,921 projects and accounting for 15.00% of the portfolio. It encompasses a broad range of social-science, humanities and policy-oriented research. Relevant projects address public policy, governance, welfare systems, inequality, gender and women's experiences, cultural identities, communities, social inclusion, education, migration, employment, law, media and cultural participation. The topic may also include research on public-service delivery, disadvantaged social groups, family and household life, political

institutions, heritage and the social consequences of economic and technological change. It captures research on how institutions, policies, and cultural practices shape social experiences and outcomes.

3.1.2 Topic 1: Climate, plants, crops and species

Topic 1 contains 12,571 projects, representing 9.96% of the portfolio. It brings together environmental, ecological, agricultural and biological research concerned with climate change and living systems. The topic includes climate modelling, adaptation and mitigation, biodiversity and species conservation, ecosystem change, plant biology, crop improvement, agricultural productivity, soil and forest systems, pests and plant diseases, and the effects of environmental stress on organisms. It may also encompass food security, sustainable agriculture, marine and freshwater ecology, ecological monitoring, land-use change and the management of natural resources. The topic reflects the close connection between climate research, environmental protection and the resilience of agricultural and ecological systems.

3.1.3 Topic 2: Cells, cancer, proteins and infection

Topic 2 contains 11,731 projects and accounts for 9.30% of the portfolio. It represents a major cluster of molecular, cellular and biomedical research. Relevant projects investigate cellular functions, protein structures and interactions, cancer development and treatment, immune responses, infectious diseases, pathogens and host–pathogen interactions. The topic also covers molecular mechanisms of disease, cell signalling, tumour biology, microbiology, virology, drug targets, diagnostic biomarkers and the development of new therapeutic approaches. Genomics, biochemical methods, laboratory models, and translational research may also appear within this topic when used to understand cancer, infection, or cellular disease processes.

3.1.4 Topic 3: Quantum technologies, devices, chemical reactions and batteries

Topic 3 contains 10,746 projects, accounting for 8.52% of the portfolio. It combines several closely related areas of physics, chemistry, materials science and engineering. The topic includes quantum physics and quantum technologies, electronic and photonic devices, semiconductors, sensors, nanomaterials, chemical reactions, catalysis, electrochemistry, batteries and energy-storage systems. Projects may examine the design and fabrication of new devices, the properties of functional materials, reaction mechanisms, charge transfer, interfaces and surfaces, and improvements in battery performance, durability and safety. The topic captures research extending from fundamental quantum and chemical processes to practical technologies in electronics, sensing, communications and low-carbon energy.

3.1.5 Topic 5: Brain, neurons, memory and neural systems

Topic 5 contains 4,715 projects, corresponding to approximately 3.74% of the portfolio. It represents research in neuroscience, cognitive science and related clinical fields. Relevant projects investigate brain structure and function, neurons and neural circuits, memory, learning, perception, cognition and behaviour. The topic also includes brain development, neuroplasticity, neuroimaging, computational neuroscience, neural networks and brain–computer interaction. Clinical applications may concern neurodegenerative disorders, neurological diseases, mental health conditions, brain injury and the development of diagnostic or therapeutic interventions. The topic connects fundamental research on neural mechanisms with research on cognition, behaviour and brain-related illness.

3.1.6 Topic 4: Galaxies, stars, the universe and physics

Topic 4 contains 4,490 projects and accounts for 3.56% of the portfolio. It represents astronomy, astrophysics, cosmology and related areas of fundamental physics. Relevant projects address the formation and evolution of stars and galaxies, planetary and stellar systems, black holes, gravitational phenomena, dark matter, dark energy and the large-scale structure and history of the universe. The topic may also include astronomical observation, telescope and detector development, space instrumentation, computational simulations and the analysis of large astronomical datasets. It combines theoretical physics, observational astronomy, scientific instrumentation and data-intensive modelling.

3.1.7 Topic 9: Security, business, services and users

Topic 9 contains 4,433 projects and represents 3.51% of the portfolio. This is a broad digital, organisational and service-oriented topic. It includes cybersecurity, information privacy, digital trust, secure networks and platforms, business information systems, financial and professional services, enterprise technologies and the digitalisation of public services. The presence of the words "services" and "users" indicates that the topic also encompasses user experience, human–computer interaction, service design, technology adoption and the relationships between digital systems and their organisational or individual users. Some projects may examine cyber risk and technical security. In contrast, others focus on how businesses, public organisations and consumers use, govern and experience digital services.

3.1.8 Topic 11: Care, patients and healthcare

Topic 11 contains 3,187 projects, accounting for 2.53% of the portfolio. It focuses on patient-facing research and the organisation and delivery of healthcare. Relevant projects address diagnosis, treatment, clinical pathways, patient safety, chronic-disease management, nursing, social care, rehabilitation, palliative care and community-based services. The topic also encompasses healthcare access, health inequalities, service quality, workforce organisation, patient experiences and the implementation of medical technologies in clinical practice. Digital health, remote monitoring, healthcare data and the evaluation of new interventions may be

included when their primary purpose is to improve patient outcomes or healthcare delivery. This topic is more directly concerned with care systems and clinical application than Topic 2, which focuses more heavily on cellular and molecular biomedical mechanisms.

3.1.9 Topic 6: Equations, mathematics and geometry

Topic 6 contains 3,165 projects and represents 2.51% of the portfolio. It covers pure and applied mathematical research. The topic includes algebra, geometry, topology, mathematical analysis, differential equations, number theory, probability, statistics, optimisation and numerical methods. It may also include mathematical physics, computational mathematics, and the construction of models used to understand complex systems in engineering, economics, biology, medicine, and the physical sciences. Some projects focus on developing abstract mathematical theories. In contrast, others apply equations and mathematical structures to scientific prediction, simulation and decision-making.

3.1.10 Topic 7: Music, games and sound

Topic 7 contains 2,892 projects, accounting for 2.29% of the portfolio. It brings together research in music, sound, games, digital media and the creative industries. Relevant projects include music composition, performance, musicology, sound art, acoustics, audio engineering, music technology, game design and game development. The topic also encompasses spatial and immersive audio, auditory perception, sound recognition, film and television sound, virtual and augmented reality, interactive media and human–computer interaction. It includes both arts-and-humanities research on music and cultural production and technical research on audio systems, digital entertainment and immersive creative experiences.

3.2 Hierarchical clustering

To better understand higher-order thematic relationships, Figure 1 shows the hierarchical clustering dendrogram of the 69 topics. The dendrogram further demonstrates that the topic structure is not random but organised into several broader semantic branches.

At the highest level, the topics can be broadly divided into two major branches. The first branch is dominated by biomedical and health-related topics, including themes associated with cell biology, cancer, infection, genes, tissues, organs, neurological conditions, chronic disease, and clinical care. These topics tend to merge at relatively short distances, indicating strong semantic relatedness and forming a coherent life and health sciences cluster.

The second major branch is broader and more heterogeneous, encompassing themes related to physical sciences, engineering, environment, social issues, and humanities-oriented topics. Within this branch, several subclusters are evident. Overall, the hierarchical clustering results suggest that the topic landscape is characterised by clear disciplinary groupings at the macro level, while still maintaining substantial internal variation within each branch. This structure reflects the corpus's multidimensional, cross-domain nature.

Hierarchical Clustering

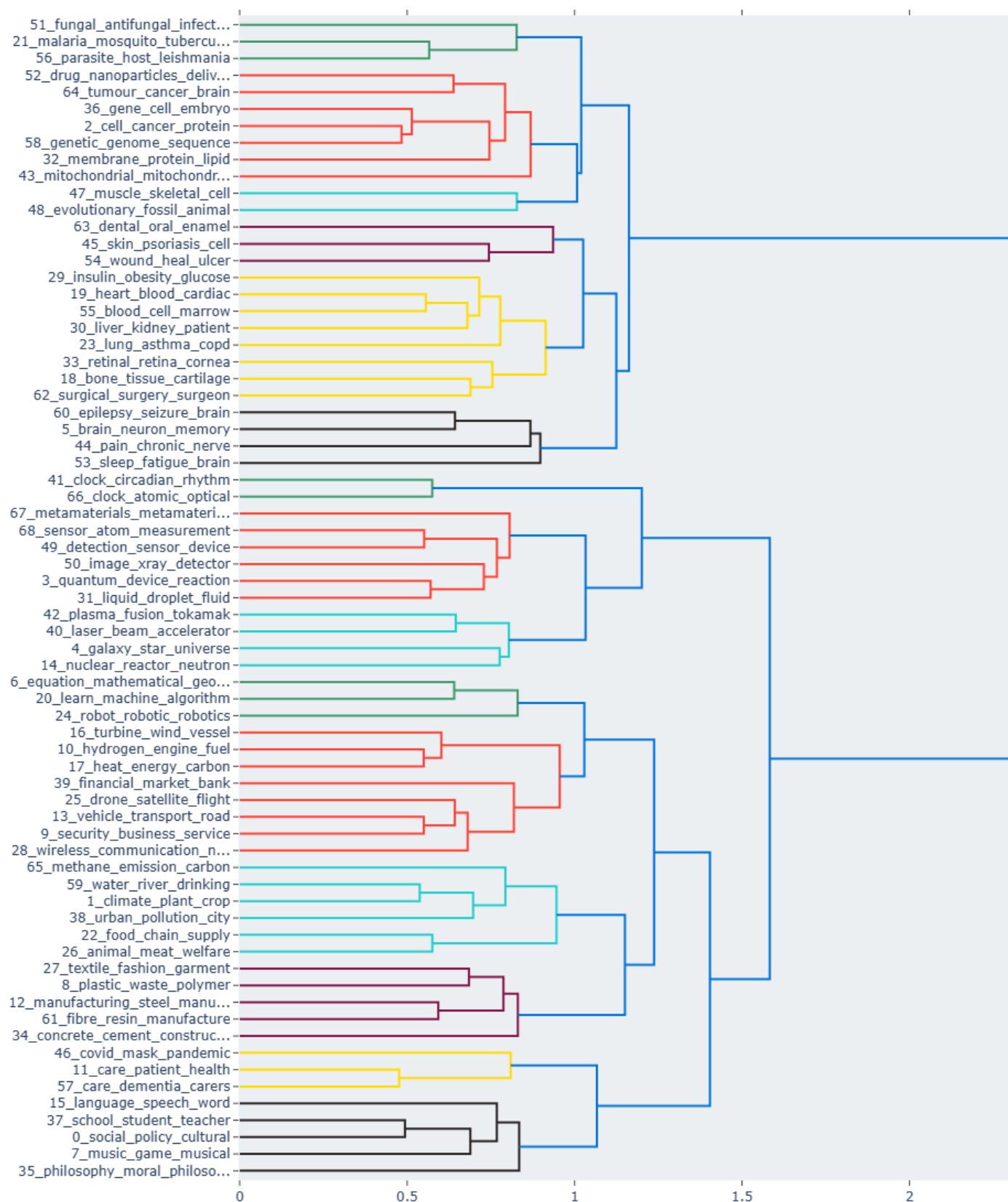


Figure 1. Hierarchical clustering

3.3 Similarity matrix

Figure 2 presents the similarity matrix of all topics. Consistent with the hierarchical clustering results, the similarity matrix shows that most topic pairs exhibit low to moderate similarity, as indicated by the predominance of lighter colours outside the diagonal. This pattern confirms that most topics are semantically distinct and that the topic model achieves a satisfactory level of differentiation. The similarity matrix suggests that the corpus exhibits a dual structure: considerable semantic heterogeneity across the full set of topics. In contrast, semantically proximate topics still cluster into broader thematic domains. This finding strengthens the interpretation that the field contains both well-defined specialised topics and higher-level integrative knowledge clusters.

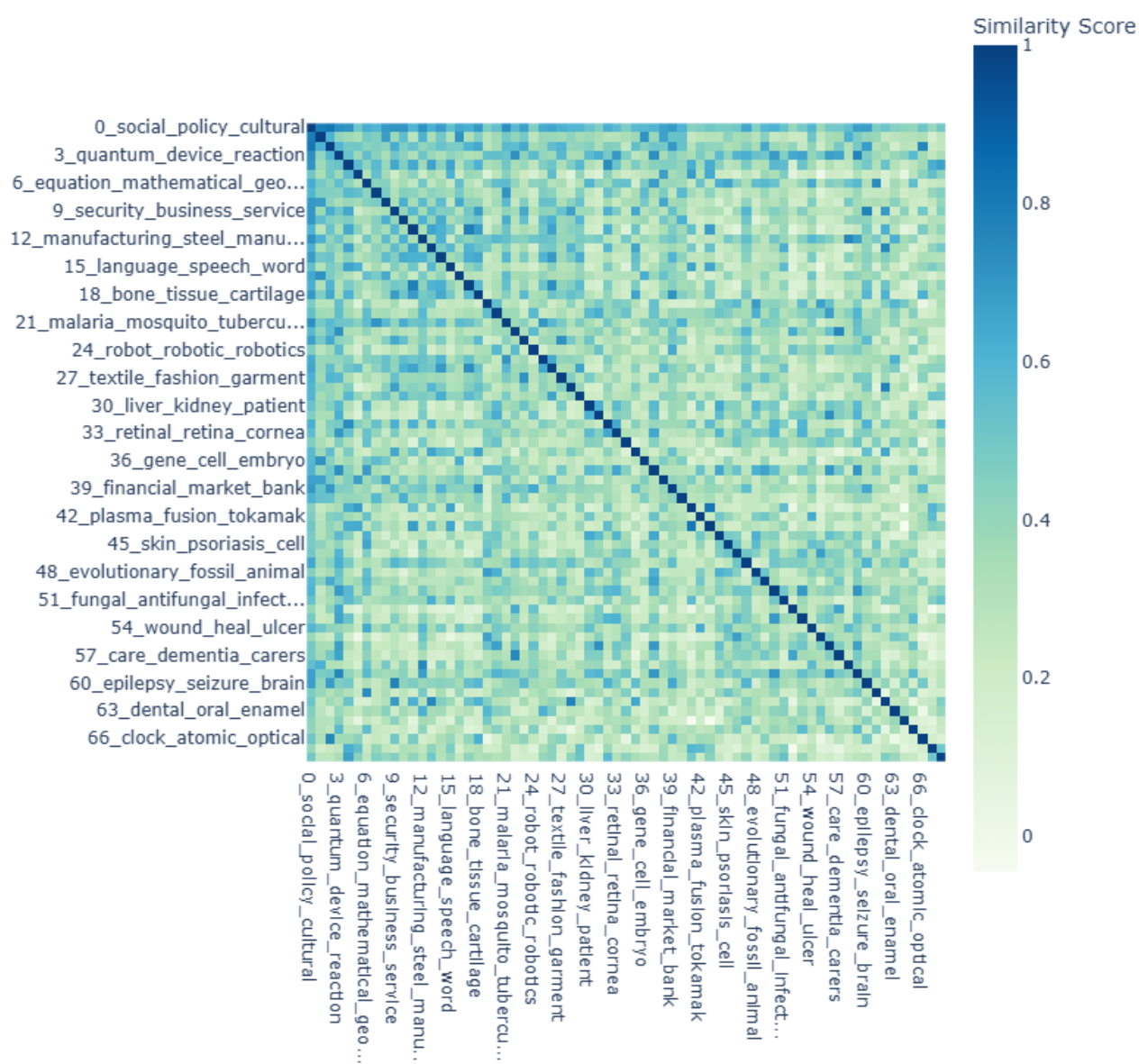


Figure 2. Similarity matrix

4. Thematic differences across UK regions

Further analysis of the results allows us to address the report's first research question: *Do UKRI-funded projects exhibit significant thematic differences across UK regions?* Before showing the regional differences, the analysis first revealed similarities in topics across different regions. First, society / policy / culture / women is the most widespread leading topic. It ranks first in 11 of the 12 regions, with regional shares ranging from 13.8% to 24.9%. Its share reached 24.9% in Northern Ireland, 21.6% in the North East, 21.3% in Wales, and 20.0% in London. Second, in 11 out of the 12 regions, the top 4 most funded projects are in some combination of society / policy / culture / women, climate / plants / crops / species, cells / cancer / proteins / infection and quantum / devices / reactions / batteries. The only exception is London, where Brain / neurons / memory / neuroscience ranked the fourth-highest topic. This regional result is consistent with the top 4 topics at the national level, reflecting national priorities across regions.

Beyond these four widely funded themes, the regional portfolios exhibited some distinctive features. These differences were generally not reflected in the largest topic in each region, which was often one of the shared nationwide themes. Instead, they appeared in the composition of the remaining top ten topics and in the concentration of particular fields within specific university cities and regional research ecosystems. We call these secondary specialisations. We examine these regional orientations below. Appendix B lists the top 10 topics that appear across regions and reflect these secondary specialisations.

4.1 East Midlands

The East Midlands region could be categorised by a group of industrial and engineering-related topics within its profile, with hydrogen / engines / fuels / vehicles, plastics / waste / polymers / products, and manufacturing / steel / metals accounting for approximately 8.4% of all projects in the region. Together, these topics collectively point to a broader regional orientation towards sustainable advanced manufacturing and industrial decarbonisation. This interpretation is consistent with the East Midlands' industrial structure. The region's longstanding industrial base in aero-engines, automotive production, rail and materials, particularly the aerospace cluster surrounding Derby, creates demand for research into low- and zero-carbon engines, hydrogen-based fuels, lightweight and recyclable materials, lower-energy production processes, industrial-plastic recycling, and improved metal and composite-manufacturing technologies. The observed thematic profile is therefore consistent with projects responding to concrete technological challenges arising from the region's existing industrial clusters.

The city-level distribution demonstrates that specific institutions mainly support this specialisation. For example, of the 190 projects classified under hydrogen / engines / fuels / vehicles, 99 were located in Nottingham and 82 in Loughborough, representing 52.1% and 43.2% of the regional total, respectively. Similar observation can be made about projects concerning plastics, waste and polymers. Of the 169 projects in this topic, Nottingham

accounted for 125(74.0%), while Loughborough accounted for 27(16.0%). Projects on manufacturing, steel and metals were similarly concentrated. Of the 142 projects assigned to this topic, Nottingham contributed 82 (57.7%), and Loughborough contributed 46 (32.4%). Loughborough and Nottingham appear to have provided important institutional anchors through their universities and research centres. For example, the National Centre for Combustion and Aerothermal Technology in Loughborough builds on a long-term partnership with Rolls-Royce and conducts research on aero-engine fuel efficiency, hydrogen-fuel technologies and sustainable aviation fuels.

Consequently, although industrial demand is distributed more broadly across the region, including the prominent manufacturing base surrounding Derby, the research projects themselves are concentrated in university-based cities with laboratories, infrastructure, specialised researchers, and the administrative capabilities required to develop and manage funded projects.

4.2 East of England

The East of England region also exhibits a modest relative specialisation in projects concerning hydrogen / engines / fuels / vehicles. By contrast, although machine learning / algorithms appeared among its ten leading topics, its regional share was not higher than that observed in the national project portfolio. As in the East Midlands, research institutions in the region support this specialisation. Of the 172 projects classified under hydrogen / engines / fuels / vehicles, Cambridge accounted for 122(70.9%), while Cranfield accounted for 44 (25.6%).

Specifically, Cranfield provides the more strongly applied and infrastructure-intensive component of this regional specialisation. The university has established capabilities in aerospace, propulsion, vehicle engineering and energy systems. It's strength in hydrogen research, supported by the £69 million hydrogen centre, encompassing hydrogen-fuelled engines, fuel cells, sustainable transport fuels, hydrogen production and storage, aircraft propulsion, as well as airport infrastructure and operations. Thereby enabling the development and testing of hydrogen technologies across the aviation supply chain.

Similarly, Cambridge contributes complementary capabilities in propulsion science, sustainable aviation, energy production and whole-system analysis. The Whittle Laboratory has a longstanding research focus on propulsion and power and has increasingly directed these capabilities towards zero-carbon aviation. Its work encompasses a range of potential propulsion pathways, including battery-electric systems for shorter-distance flight and hydrogen-powered aircraft for longer-distance applications. In addition, Cambridge participates in the Global Hydrogen Production Technologies Centre, a five-year initiative co-led by Cranfield that seeks to improve the technological and economic viability of net-zero hydrogen production. The combination of Cranfield's applied aviation and hydrogen infrastructure with Cambridge's capabilities in propulsion, hydrogen production and system-level aviation

decarbonisation has consequently produced a recognisable regional cluster in hydrogen-enabled transport and low-carbon aviation.

4.3 London

London exhibited a distinctive thematic orientation towards patient care and healthcare. This is supported by substantial but less regionally concentrated capabilities in machine learning and digitally enabled services. Projects classified under care / patients / healthcare accounted for 2.52% of London's portfolio, compared with 1.82% in the full sample. This was the highest share among the 12 regions, making healthcare a recognisable secondary specialisation within London's funded-project portfolio.

This concentration is consistent with London's dense health and life-sciences ecosystem. The capital brings together major universities and medical schools, teaching hospitals, academic health-science partnerships, NHS organisations, medical research charities, and pharmaceutical, biotechnology and medical-technology firms. The Greater London Authority accordingly characterises London as a global life-sciences hub whose network spans universities, research centres, healthcare providers, medical charities, innovative firms and major industry actors, with activities extending from pharmaceuticals and advanced therapies to medical technologies and healthcare AI. This institutional density provides both the research capabilities and the practical environments needed to conduct projects directly concerned with patients and healthcare provision. Relevant research projects concerned may address patient diagnosis and treatment, long-term and chronic-disease management, the organisation and delivery of healthcare services, inequalities in access to care, the clinical application of medical technologies, and the use of patient data to improve treatment and service design. These research activities shows linkages and extension of fundamental biomedical research into patient care, clinical translation and the improvement of healthcare systems.

Machine learning and algorithms also constituted an important element of London's project portfolio, accounting for 2.16% of projects, compared with 1.94% in the full sample. London possessed a substantial machine-learning portfolio whose scale reflected the breadth of its universities, research institutes and technology sector. The wider London AI ecosystem provides an important institutional context for this portfolio. The King's Cross area, for example, has developed into a substantial AI cluster containing major international and British AI companies, including Google DeepMind, Anthropic, Meta, Perplexity, Scale AI, Wayve and Synthesia. More generally, UK evidence shows that AI investment extends well beyond information technology into life sciences and biotechnology, automotive and transport, financial services, and healthcare and wellness. London's machine-learning projects are consequently likely to span medical diagnosis and patient-risk prediction, drug discovery and medical imaging, financial-risk analysis and fraud detection, cybersecurity and privacy, public-service management, business analytics, consumer behaviour, language, media and the creative industries.

As a result, London's most distinctive orientation therefore lies in its intersection: AI-enabled health and service innovation supported by expertise in data security, privacy and user-centred digital systems. Healthcare is the clearest regionally overrepresented theme, while machine learning and secure digital services operate primarily as complementary capabilities connecting London's medical, financial, professional-service, and public-sector ecosystems.

4.4 North East

The North East exhibits a distinctive thematic profile that combines fundamental science with secure and data-enabled services. These orientations are visible in the region's secondary topics, beyond the four themes widely funded across the United Kingdom. A prominent feature of the North East portfolio is its concentration in fundamental science. Projects concerning galaxies / stars / universe / physics account for 5.8% of the regional portfolio, while equations / mathematics / geometry account for a further 3.0%. The astronomy component gives the region a particularly visible basic-science identity. This profile is strongly anchored in Durham University, whose astronomy research cluster encompasses the Centre for Extragalactic Astronomy, the Institute for Computational Cosmology and the Centre for Advanced Instrumentation. Durham describes its Centre for Extragalactic Astronomy as one of the largest astronomy groups in the United Kingdom and Europe. At the same time, the Ogden Centre provides dedicated infrastructure for research on the universe, dark matter and dark energy. The coexistence of astronomy and mathematics also reflects the computational and quantitative requirements of contemporary astrophysics, including modelling galaxy formation, cosmological simulation, statistical inference, and the analysis of large observational datasets.

A second regional orientation concerns security / business / services / user. This topic accounts for 2.22% of North East projects, giving the region the second-highest share among the 12 regions. Its 99 projects are concentrated primarily in Newcastle upon Tyne, which contributes 59 projects (59.6%), and Durham, which contributes 23 (23.2%). Middlesbrough and Sunderland contribute a further 12 and 5 projects, respectively. The topic extends beyond cybersecurity alone. It includes cybersecurity and digital resilience, enterprise software and information systems, financial technology, secure digital platforms, data privacy, user trust, public-service digitalisation and user-centred service design. These areas correspond closely to the North East's established capabilities in software development, programming, service centres, financial technology, gaming and cyber-threat management. Regional evidence identifies particular strengths in software, data analytics, platforms and cybersecurity, alongside a wider digital ecosystem containing companies such as Sage and a substantial concentration of specialist technology businesses. Newcastle University provides a major institutional anchor for this orientation through the National Innovation Centre for Data. Created through investment from Newcastle University and the UK Government, the centre works with private companies, public-sector organisations and charities to apply data science and artificial intelligence to practical organisational problems, develop new products and services, and strengthen data capabilities in the workforce. The regional topic can therefore be characterised

as secure and data-enabled business and public services, supported by university expertise and a growing software, financial-technology and digital-service economy.

4.5 North West

The North West exhibits a unique specialisation, combining astronomy and nuclear in its profile. Projects concerning galaxies / stars / universe / physics account for 6.5% of the regional portfolio, while nuclear energy / reactors / neutrons / radioactivity account for a further 4.1%. Specifically, the astronomy speciality reflects in the North West's established research institutions and large-scale scientific infrastructure. The University of Manchester's Jodrell Bank Centre for Astrophysics is one of the United Kingdom's largest astronomy research groups. Its research spans radio astronomy, stars, pulsars, galaxies and cosmology.

Meanwhile, its facilities include e-MERLIN, the United Kingdom's national radio-astronomy and very-long-baseline-interferometry facility, as well as high-performance computing and astronomical data-processing capabilities. Liverpool also contributes through the Astrophysics Research Institute at Liverpool John Moores University, particularly in observational and theoretical astrophysics. These institutions support a regionally concentrated capacity in astronomical observation, instrumentation, computation and the analysis of large scientific datasets.

In addition, nuclear research is the North West's strongest relative specialisation. Of the 363 nuclear-related projects in the North West, Manchester accounts for 234 (64.5%), Liverpool for 88(24.2%), and Lancaster for 37(10.2%). The project bases are therefore concentrated in major university cities, while the wider regional specialisation draws heavily on West Cumbria's nuclear-industrial and research infrastructure. Here, Sellafield provides a major source of technological capabilities through its activities in nuclear decommissioning, radioactive-waste processing, environmental restoration, fuel and material management, and the handling of hazardous radioactive environments. In addition, the UK National Nuclear Laboratory (UKNNL) operates its Central Laboratory and Windscale Laboratory at Sellafield. The Central Laboratory supports reactor operations, new reactor development, decommissioning, clean-up, and fuel processing, while the Windscale Laboratory examines irradiated fuel and materials, radioactive-waste processing, and material testing. UKNNL also maintains facilities in Preston and Workington, strengthening the wider North West nuclear research infrastructure. The concentration of nuclear projects reflects a sustained interaction among large-scale decommissioning and waste-management tasks, specialised nuclear laboratories, university research capacity and long-term industry–university collaboration. This gives the Northwest a distinctive profile across the nuclear lifecycle, from reactor and fuel research to decommissioning, environmental restoration, and radioactive-waste management.

4.6 Northern Ireland

Northern Ireland exhibits strong regional specialisations concerning security / business / services / users, lasers / beams / accelerators / electrons, and wireless communications / networks / optics. In addition, projects on music / games / sound provide an additional creative-technology orientation. First, projects classified under security / business / services / users account for 3.68% of the Northern Ireland portfolio. Of the 53 projects in this topic, Belfast contributes 43 (81.1%), while Coleraine contributes 10 (18.9%). This specialisation is closely connected to the Centre for Secure Information Technologies at Queen's University Belfast. Another representative technology-related topic is lasers / beams / accelerators / electrons. It accounts for 2.71% of Northern Ireland's projects, compared with approximately 0.42% in the full sample. All 39 projects in this topic are located in Belfast, demonstrating an exceptionally concentrated institutional specialisation. This profile reflects Queen's University Belfast's long-term capabilities in high-intensity laser and plasma physics. Its Centre for Plasma Physics has developed research on laser- and electrically generated plasmas, femtosecond and high-intensity lasers, and interactions among plasmas, charged particles and matter. The third major specialisation concerns wireless communications / networks / optics. The topic accounts for 1.95% of Northern Ireland's projects and has the highest regional share in the United Kingdom. Belfast contributes 25 of the 28 relevant projects (879.3%), while Coleraine contributes three. This concentration corresponds directly with the Centre for Wireless Innovation at Queen's University Belfast. Queen's concentration of expertise in electronics, wireless systems and secure information technologies enables projects to address communications performance and security within the same institutional environment. This connection gives Northern Ireland a broader specialisation in secure and advanced connected systems.

In addition to the technological orientation, projects classified under music / games / sound contribute a complementary creative and digital orientation. The topic accounts for 2.50% of the Northern Ireland's portfolio, compared with 1.68% in the full sample, giving the region the second-highest share among the 12 regions. Belfast contributes 31 of the 36 relevant projects (86.1%), while Coleraine contributes five (13.9%). The Belfast concentration is closely associated with the Sonic Arts Research Centre (SARC) at Queen's University Belfast. SARC was established as an interdisciplinary research environment at the intersection of music, technology, sound and society. Its Sonic Laboratory supports research into virtual and augmented reality, immersive media, multichannel and spatial audio, sound engineering, musical interaction, instrument design and music perception. Ulster University adds capabilities in composition, music technology, performance, soundtrack research, computer games and creative industries. The regional topic therefore encompasses music and sound technologies, immersive media and creative digital production. It connects artistic practice with audio engineering, human-computer interaction, virtual environments and digital-media development. These connections also create potential overlaps with wireless communications, artificial intelligence and user-centred technology.

4.7 Scotland

Scotland exhibits two complementary thematic orientations within its UKRI-funded project portfolio. The first and most distinctive concerns offshore wind, turbine technologies, marine engineering and offshore energy operations. The second is an Edinburgh-centred concentration of capabilities in artificial intelligence, language and speech technologies. Within these two clusters, turbines / wind energy / vessels / energy provides the strongest evidence of regional specialisation, while language / speech / linguistics constitutes a more moderate but recognisable secondary orientation. Machine learning / algorithms represents a substantial cross-cutting capability that supports multiple fields across the Scottish research system.

Specifically, projects classified under turbines / wind energy / vessels / energy account for 1.96% of Scotland's portfolio, approximately 2.5 times the share observed in the full sample. Scotland has the highest regional share of this topic among the 12 regions, making it one of the clearest indicators of thematic differentiation in the Scottish project portfolio. Of Scotland's 239 projects in this topic, Edinburgh contributes 104 (43.5%), while Glasgow contributes 100 (41.8%). Aberdeen contributes a further 14 projects, or 5.9%. Edinburgh and Glasgow therefore account for approximately 85.4% of the regional total. The city distribution shows that the research projects are concentrated primarily in Scotland's two largest university centres, while the technological problems, industrial partners and application environments extend across the country's wider offshore-energy economy. The regional concentration reflects Scotland's extensive offshore wind resources and its growing pipeline of fixed and floating offshore developments. Programmes such as ScotWind and INTOG create sustained research demand in wind-turbine design, floating foundations, offshore structures, power transmission, installation, operations and maintenance, and supply-chain development. Scotland's offshore-wind ambitions further reinforce the long-term importance of research that can reduce costs, improve reliability and support the deployment of wind-energy systems in increasingly complex marine environments. The Offshore Renewable Energy Catapult operates wind-energy demonstration and research facilities in Scotland, including turbine and floating-wind capabilities. At the same time, the University of Strathclyde contributes expertise in turbine engineering, marine structures, offshore installation, logistics and operations and maintenance.

Projects on machine learning / algorithms form another substantial component of Scotland's portfolio. They account for 1.88% of Scottish projects, close to the 1.94% share in the full sample. Of the 229 projects in this topic, Edinburgh contributes 140, or 61.1%; Glasgow contributes 59, or 25.8%; and Aberdeen contributes 11, or 4.8%. Edinburgh and Glasgow together account for approximately 86.9% of the regional total. The concentration reflects Edinburgh's longstanding artificial-intelligence research ecosystem, supported by large academic communities in informatics, computer science, data science and related fields. Institutions such as the Bayes Centre connect university researchers, students, start-up firms and established companies and facilitate the application of AI and data science to industrial and societal challenges. Glasgow adds further capabilities in machine learning, computer

vision, robotics, autonomous systems and data-intensive science. Machine learning supports a wide range of Scottish research activities.

The third theme, language / speech / linguistics, accounts for 1.83% of Scotland's portfolio, compared with 1.50% in the full sample. Scotland has the third-highest regional share, following Wales and the North West. This pattern constitutes a moderate regional specialisation, strongly concentrated in Edinburgh and Glasgow. Of the 223 relevant projects, Edinburgh contributes 124 (55.6%), while Glasgow contributes 53 (23.8%). Aberdeen contributes a further 17 projects (7.6%). Edinburgh and Glasgow together account for approximately 79.4% of the regional total, with Edinburgh serving as the principal research node. The University of Edinburgh has longstanding capabilities in speech and language technology. Its Centre for Speech Technology Research focuses on speech recognition, speech synthesis, speech-signal processing, dialogue systems and multimodal interaction. Edinburgh's natural-language-processing research also encompasses language models, syntax, semantics, text generation, machine translation and computational approaches to human communication. These activities connect researchers in informatics, linguistics, psychology and cognitive science.

4.8 South East

The South East represents a digital and computational research profile centred on machine learning and its application across science, medicine, engineering and interactive media. Projects concerning machine learning / algorithms provide the clearest evidence of regional specialisation. In addition, music / games / sound constitutes a moderate but institutionally grounded creative-technology orientation.

First, projects classified under machine learning / algorithms account for 2.50% of the South East portfolio, giving the region the second-highest share among the 12 regions, following the South West. The geographical distribution is strongly concentrated in two major academic centres, namely Oxford and Southampton. Oxford provides the principal institutional anchor for this specialisation. Its machine-learning community encompasses both foundational methodological research and applications across multiple scientific and technological fields. The Oxford Applied and Theoretical Machine Learning group conducts research within the Department of Computer Science, while the wider Oxford AI community includes expertise in statistical learning, computer vision, robotics, biomedical imaging, digital health, autonomous systems, materials discovery and AI for science and engineering. This breadth helps explain why machine learning appears across projects in medicine and the life sciences, autonomous vehicles and robotics, astronomy, image and signal processing, engineering systems, risk analysis and scientific discovery. Southampton provides a second large research node. Its School of Electronics and Computer Science brings together more than 550 researchers working across artificial intelligence, machine learning, Web Science, electronics and related digital fields. AI@Southampton connects these capabilities across the university, facilitating interdisciplinary research, education and knowledge exchange in artificial intelligence. Southampton's contribution therefore extends beyond conventional computer science into

robotics, autonomous systems, engineering, healthcare technologies, public services and digital society. This institutional concentration provides the region with a broad computational foundation that can support research across its scientific, medical, engineering and digital-service ecosystems.

In addition, the South East has the fourth-highest share among the 12 regions in Projects concerning music / games / sound. In terms of cities, Southampton contributes a longstanding concentration in acoustics, sound engineering and auditory technologies through its Institute of Sound and Vibration Research. ISVR research examines how sound is generated, transmitted, captured, manipulated, controlled and perceived, while specific projects cover 3D audio recording and reproduction, loudspeaker and microphone arrays, sound-field control, signal processing and immersive listening technologies. Southampton's projects in this topic are consequently likely to include acoustics, audio engineering, spatial audio, hearing and audiology, sound recognition, and the control of noise and vibration. Guildford provides a different component based on video games, immersive media and creative computing. The University of Surrey describes Guildford as the centre of one of the United Kingdom's largest video-game clusters. Its games-related capabilities connect coding, animation, motion capture, game AI, sound design and interactive production. Surrey's Creative Vision and Sound research further combine machine perception, performance capture, immersive virtual-reality production, spatial audio, machine audition and sound recognition for film and games. Guildford's contribution can therefore be characterised as the intersection of games, artificial intelligence, immersive media and spatial sound. Brighton adds a broader creative-industries and digital-culture environment. Greater Brighton and Coastal West Sussex contain an extensive creative economy covering digital media, games, music, design, performance and related services. A regional creative-industries assessment identifies 10,225 companies and a workforce of 54,428 people within this ecosystem. This industrial setting creates opportunities for university collaboration and applied research concerning creative production, audiences, user experience, digital platforms and new forms of cultural participation.

4.9 South West

Similar to the South East, the South East region demonstrated an orientation towards machine learning / algorithms, combined with Manufacturing / steel / metals to add an established advanced-manufacturing and materials component. Specifically, projects classified under machine learning / algorithms are particularly prominent in the South West. The region contains 199 projects in this topic: Bristol contributes 124 (62.3%), Bath contributes 37 (18.6%), and Exeter contributes 30 (15.1%). The regional specialisation is therefore anchored in a Bristol–Bath–Exeter research network with substantial capabilities in artificial intelligence, data science, robotics and intelligent systems. Bristol has a longstanding research base in artificial intelligence, machine learning, robotics and autonomous systems. Bath contributes interdisciplinary AI capabilities that connect machine learning with engineering, the life sciences, chemistry and physics. Exeter's Institute for Data Science and Artificial Intelligence

similarly applies AI to environmental, health and societal challenges. More recent infrastructure, including the Isambard-AI national computing facility located at the National Composites Centre, demonstrates the continued expansion of this regional capability. Because this infrastructure may postdate some projects in the dataset, it is best treated as evidence of the longer-term accumulation and reinforcement of existing research strengths.

Projects concerning manufacturing / steel / metals constitute a second component of the regional profile. The South West contains 172 projects in this topic. Bristol contributes 119, or 69.2%; Bath contributes 38, or 22.1%; and Exeter contributes 14, or 8.1%. Research in this area is therefore concentrated overwhelmingly in the Bristol–Bath advanced-engineering system. The South West, particularly the West of England, contains one of the United Kingdom's most substantial aerospace and advanced-engineering clusters. The regional ecosystem includes major industrial organisations such as Airbus, Rolls-Royce and GKN, alongside research institutions such as the National Composites Centre and the Bristol Composites Institute. These organisations create sustained research demand for lighter, stronger aerospace components; faster, more precise manufacturing; reduced energy and material consumption; recyclable or lower-impact composites; and the use of digital technologies to optimise design, production and operation. The National Composites Centre and the Bristol Composites Institute strengthen the complete research chain extending from materials design and manufacturing experimentation to testing and industrial implementation. The regional capability can therefore be characterised as advanced materials, aerospace manufacturing and digitally enabled production.

4.10 Wales

Wales exhibits two highly distinctive thematic orientations within its UKRI-funded project portfolio. The first concerns the green transformation of a materials-intensive industrial economy, expressed through projects on manufacturing, steel, metals, plastics, waste and polymers. The second concerns the preservation and digital development of the Welsh language, combining linguistic research, bilingual public policy and language technologies. Among the three themes, manufacturing / steel / metals and language / speech / linguistics provide the strongest evidence of regional specialisation.

Projects classified under manufacturing / steel / metals constitute Wales's clearest industry-related specialisation. Wales contains 128 projects in this topic, of which Swansea contributes 89 (69.5%) and Cardiff contributes 37 (28.9%). Together, the two cities account for more than 98% of the regional total, with Swansea serving as the principal research node. This spatial concentration reflects the interaction between South Wales's longstanding steel and metal-processing base and Swansea's specialised research capabilities. South Wales has a substantial industrial history in steel, metals and advanced manufacturing, particularly around Port Talbot.

The wider Welsh high-value manufacturing economy also encompasses advanced materials, automotive production, aerospace and photonics. These industries create sustained research demand concerning material performance, production technologies, process efficiency and industrial decarbonisation. Swansea University's Steel and Metals Institute provides an important institutional connection between these industrial challenges and the funded-project portfolio. Developed in collaboration with Tata Steel, the institute supports research on alloy development, process simulation, mechanical testing, material characterisation and the decarbonisation of steel production. It also connects university researchers with industrial engineers and firms across the steel and manufacturing supply chain. Swansea has also led or participated in the SUSTAIN Future Manufacturing Research Hub, which addresses carbon-neutral steel production, resource efficiency and intelligent steel processing. The manufacturing / steel / metals topic consequently encompasses a wider set of activities than conventional steel production. It includes low-carbon steelmaking, industrial decarbonisation, advanced alloys and materials, metal surfaces and coatings, intelligent manufacturing, process control, material circularity, and applications in the aerospace and automotive sectors.

A second highly distinctive Welsh specialisation concerns language / speech / linguistics. This topic accounts for 2.46% of the Welsh project portfolio and has the highest regional share among the 12 regions. Of the 88 relevant projects, Cardiff contributes 45 (51.1%); Bangor contributes 23 (26.1%); and Swansea contributes 16 (18.2%). The distribution reveals a multi-city research system in which Cardiff, Bangor and Swansea provide complementary capabilities in linguistics, bilingualism, speech research and language technology. The prominence of this theme reflects the distinctive social, cultural and policy position of the Welsh language. Welsh is simultaneously an object of academic research, a medium of education and public service, an element of cultural continuity, and a language whose continued use increasingly depends on its presence in digital technologies. The Welsh Government's Cymraeg 2050 strategy aims to achieve one million Welsh speakers by 2050 while also expanding the opportunities and social environments in which the language can be used. This policy context creates research demand concerning Welsh-language acquisition, bilingual education, Welsh–English bilingualism, dialects and language variation, sociolinguistic attitudes, minority-language maintenance and the delivery of bilingual public services. Specifically, Bangor University's Language Technologies Unit has developed long-term expertise in text and speech technologies for Welsh and other Celtic languages. Its work includes spelling and grammar tools, electronic dictionaries, text-to-speech systems, speech recognition and machine translation. Cardiff contributes complementary capabilities in sociolinguistics, bilingualism and language-resource development. CorCenCC, for example, created a national corpus of contemporary Welsh incorporating written, spoken and digitally produced language. This infrastructure provides a foundation for both linguistic analysis and the development of Welsh-language technologies. Wales's specialisation in this topic therefore combines Welsh-language preservation, bilingualism, speech technology, natural-language processing and digital inclusion. It is strongly grounded in the region's own linguistic environment and public-policy needs. The interaction between cultural preservation and

technological development gives this orientation a particularly place-specific character within the UK research system.

4.11 West Midlands

The West Midlands exhibits a thematic profile centred on advanced manufacturing and its increasing integration with healthcare innovation. Accordingly, manufacturing / steel / metals provides the clearest evidence of regional specialisation. Care / patients / healthcare is another recognisable strength, supported particularly by Birmingham's clinical and translational research system. First, projects classified under manufacturing / steel / metals account for 2.65% of the West Midlands portfolio, the third-highest share among the 12 regions, providing clear evidence of an industry-oriented secondary specialisation. Of the 165 projects in this topic, Coventry contributes 82 (49.7%), while Birmingham contributes 75 (45.5%). This concentration reflects the combined influence of the Coventry–Warwick manufacturing research system and Birmingham's engineering capabilities. The topic covers a wider technological field than conventional steel production. It includes automotive and transport-equipment manufacturing, aerospace engineering, metals and advanced alloys, additive manufacturing, surface engineering, industrial robotics, automation, digital production, process optimisation, battery technologies and low-carbon mobility supply chains. These activities align closely with the West Midlands' longstanding industrial strengths in automotive production, rail, aerospace, metals, and transport equipment. The region's university–industry research infrastructure enables these industrial challenges to be translated into funded projects. Warwick Manufacturing Group maintains extensive partnerships with manufacturing firms. It researches materials, automotive technologies, batteries, digital systems, supply chains and future mobility. Its National Automotive Innovation Centre brings together the University of Warwick, Jaguar Land Rover, and Tata Motors to develop technologies related to electrification, autonomous vehicles, and future transport systems. Coventry University's Institute for Advanced Manufacturing and Engineering contributes a factory-oriented model of university–industry collaboration supported by facilities for robotics, automation, laser processing, testing and digital manufacturing.

Projects concerning care / patients / healthcare provide a second identifiable regional strength. The topic accounts for 2.19% of the West Midlands portfolio, compared with 1.82% in the full sample, giving the region the second-highest share among the 12 regions after London. Of the 136 relevant projects, Birmingham contributes 81 (59.6%), while Coventry contributes 45 (33.1%). Compared with the manufacturing theme, the healthcare portfolio is more clearly centred on Birmingham. Birmingham combines university-based medical and health research with several large NHS institutions, including University Hospitals Birmingham and Birmingham Women's and Children's Hospital. It also provides access to large and diverse patient populations, clinical data, translational medicine facilities, clinical-trial infrastructure, and networks focused on diagnostics, medical devices, and digital health. Birmingham Health Partners connects the University of Birmingham, NHS organisations, and other research

partners to translate scientific discoveries into new diagnostics, medicines, medical technologies, and improvements in healthcare delivery. This environment supports research on patient treatment, disease management, clinical pathways, healthcare organisation, clinical trials, medical devices, patient safety, community and long-term care, health inequalities and the adoption of new technologies within the NHS. Birmingham's clinical research facilities also support experimental and early-phase studies involving both adult and paediatric patients. The region's healthcare orientation is therefore grounded in a connected translational pathway extending from university research to clinical testing, NHS implementation and improvements in patient care.

4.12 Yorkshire and the Humber

The Yorkshire and the Humber region showed three clearly identifiable thematic orientations within its UKRI-funded project portfolio: advanced manufacturing and metals, music and sound technologies, and nuclear science and engineering. First, Projects classified under manufacturing / steel / metals provide one of the clearest industry-related orientations. Of the 233 relevant projects, Sheffield contributes 149(63.9%); Leeds contributes 56(24.0%); and Huddersfield contributes 19(8.2%). The strong concentration in Sheffield connects the regional portfolio directly to South Yorkshire's longstanding industrial base in steel, metalworking, forging, casting and high-value engineering. This industrial heritage has developed into a research-intensive advanced-manufacturing ecosystem. The Advanced Manufacturing Innovation District brings together the University of Sheffield Advanced Manufacturing Research Centre, major industrial organisations and specialist production facilities. The AMRC conducts research in machining, materials, automation, robotics, digital manufacturing, and industrial implementation, working with companies in aerospace, automotive production, nuclear energy, and other high-value manufacturing sectors. The manufacturing / steel / metals topic therefore encompasses a broad range of activities, including advanced alloys, metal processing, aerospace and automotive components, forging, welding, precision engineering, additive manufacturing, robotics, automation and digitally enabled production.

Second, projects classified under music / games / sound form a second highly distinctive orientation. The topic accounts for 2.57% of the regional portfolio, giving Yorkshire and the Humber the highest share among the 12 regions. Of the 223 relevant projects, York contributes 102(45.7%); Leeds contributes 54,(24.2%); Sheffield contributes 39(17.5%); and Huddersfield contributes 22 (9.9%). In this vein, York contributes substantial capabilities in audio engineering, spatial sound and immersive media. The University of York's AudioLab researches spatial and immersive audio, virtual production, virtual and extended reality, digital signal processing, acoustics, psychoacoustics, music-information analysis and interactive sound systems. This research supports applications in film, television, games, accessibility, environmental sound and digital media. Huddersfield adds a complementary orientation through the Centre for Research in New Music, which has developed long-term expertise in contemporary composition, sound art, music technology and experimental creative practice.

Its research extends into film, visual media, game sound, electroacoustic music and multimedia production. Leeds and Sheffield contribute additional capabilities in music, media, cultural production and digital creative industries.

Third, projects concerning nuclear energy / reactors / neutrons / radioactivity constitute a third major regional specialisation. The topic accounts for 2.40% of the project portfolio, approximately twice the share in the full sample, giving Yorkshire and the Humber the second-highest regional concentration after the Northwest. Of the 209 projects in this topic, Sheffield contributes 84 (40.2%); York contributes 68 (32.5%); Leeds contributes 35 (16.7%); and Huddersfield contributes 22 (10.5%). Specifically, Sheffield provides the strongest applied and manufacturing-oriented component. York contributes a more fundamental nuclear-science capability. Research at the University of York addresses nuclear structure, nuclear reactions, isotopes, nuclear astrophysics and the theoretical and experimental study of nuclear processes. This gives the regional portfolio a basic-science dimension extending beyond the manufacture and operation of nuclear technologies. Huddersfield adds accelerator science and nuclear-materials research through the International Institute for Accelerator Applications. Relevant work concerns particle accelerators, particle–material interactions, radiation effects, fusion materials and the application of accelerator technologies in energy, medicine and industry. Leeds contributes further projects across nuclear materials, engineering, physics and related scientific fields.

5. Evolution of thematic heterogeneity

The analysis of thematic heterogeneity across UK regions has revealed insights into secondary specialisation. Based on topic ranks across regions, we further analysed whether secondary specialisation has changed over time, addressing our second research question. Accordingly, Figure 3 first presents the overall changes in topic popularity in the UK.

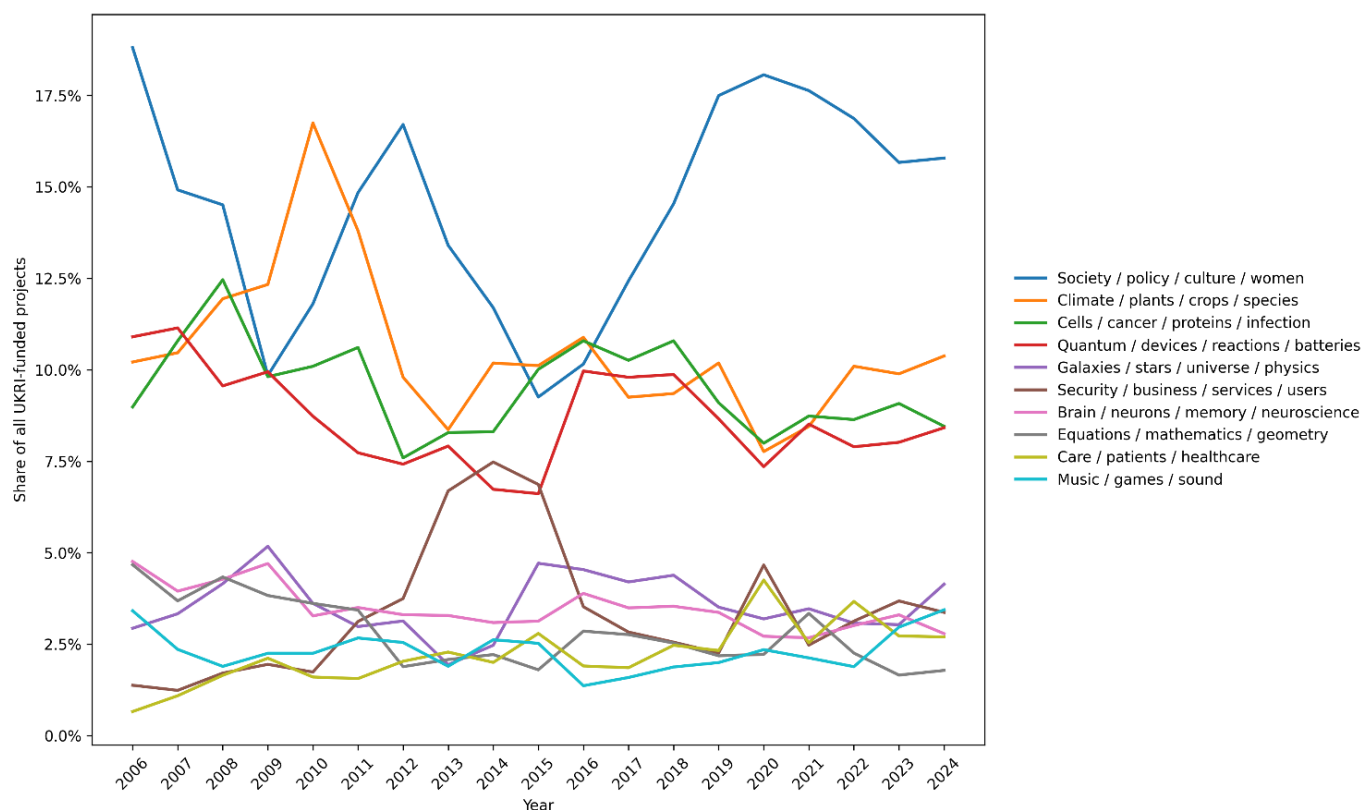


Figure 3. UK overall annual topic shares, top 10 topics (2006–2024)

At the UK level, the thematic composition of UKRI-funded projects remained broadly stable over the study period. Although annual topic shares fluctuated and some themes gained or lost relative prominence at different times, these movements were generally moderate. They did not substantially alter the overall structure of the research portfolio. The same broad group of major themes continued to account for a large proportion of funded projects throughout 2006–2024, suggesting considerable persistence in the national knowledge base. Rather than showing the displacement of established research areas by an entirely new set of topics, the temporal pattern is characterised mainly by incremental adjustment within an otherwise stable thematic structure.

This continuity is also evident in the overall level of topic concentration. The combined share of the ten largest topics declined only moderately, from approximately 64.4% in 2006–2010 to 60.4% in 2021–2024. Thus, while the portfolio became somewhat more dispersed over time, the magnitude of this change was relatively limited. A substantial majority of funded projects remained concentrated within the same broad set of research domains, and no single period

showed a fundamental reorganisation of the national thematic landscape. Increases in the relative importance of some topics were largely accompanied by gradual reductions in others, producing a process of portfolio rebalancing rather than wholesale thematic transformation.

The evidence therefore points to strong continuity with modest diversification in the UK research portfolio. The thematic landscape has evolved mainly through changes in the relative weight of established and emerging areas, rather than through major changes in its underlying composition. This suggests that the UK research system exhibits considerable thematic persistence over time: new areas can expand within the portfolio, but they tend to complement or gradually rebalance existing strengths rather than rapidly replace them. From this perspective, the most notable temporal feature is not dramatic thematic change, but the resilience of the broader national research structure despite shifts in policy priorities, technological opportunities and societal challenges over the period.

In terms of changes in different regions, however, we did not find significant differences that can be categorised as an 'evolution'. This could be anticipated, as we previously discussed; among UK regions, the top 4 funded topics are relatively consistent according to the national priorities. As place-based innovation strategy is still gaining importance in the national innovation strategy, it may still take a few years before the actual effect can be observed in the topics categorising regional funded projects. The full results of the regional-level analysis can be found in Appendix C.

6. Discussion: Identifying Six Innovation Clusters Across the UK

The results of the analysis point to two findings worth noting. First, national-level priorities are clearly reflected across regions, remaining at the top of the list despite slight differences in priorities. Although these top topics fluctuate in funded projects, changes in regional features are not significant. Second, the results still indicate regional differences in their innovation capabilities beyond the national priorities, which we call secondary specialisations. Some regions demonstrate a unique combination of such innovation capabilities; here, we further discuss how six regions are formulating clear innovation clusters with their funded projects, building up their capabilities. These six clusters are summarised in Figure 4.

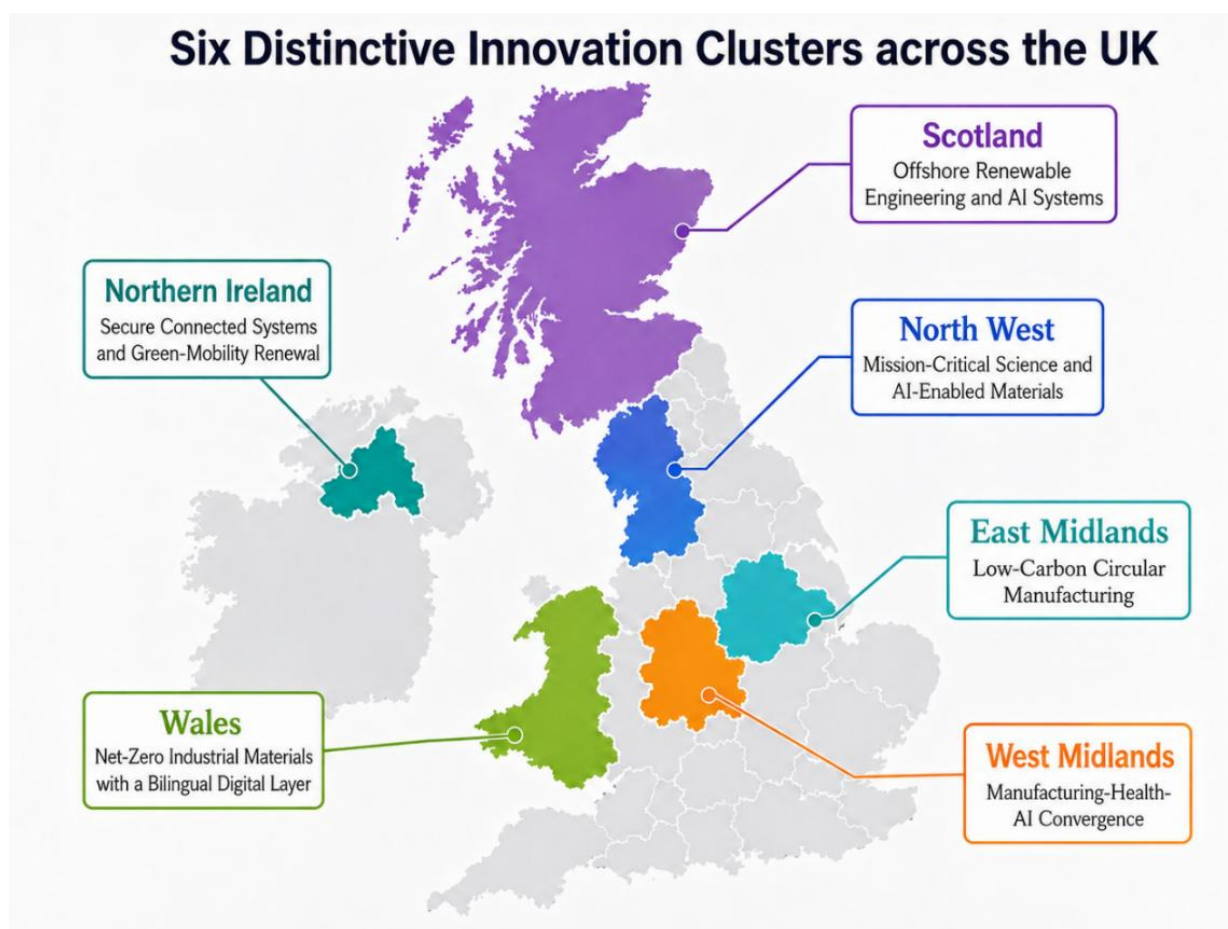


Figure 4. The six distinct innovation clusters in the UK

6.1 East Midlands - Low-Carbon Circular Manufacturing Regional innovation cluster

The East Midlands cluster is more precisely characterised as a low-carbon circular manufacturing system than as a broadly defined advanced-manufacturing cluster. Its distinctiveness lies in the way established production capabilities are being recombined with industrial decarbonisation technologies and the circular use of materials. The regional coexistence of manufacturing, plastics, hydrogen and a substantial climate-research base

creates a coherent problem–technology–application configuration. Rather than representing an entirely new industrial path, this configuration reflects related diversification around an inherited manufacturing core (Balland et al., 2019).

Within this system, Nottingham and Loughborough form a concentrated knowledge-generation and translation axis, while the region’s manufacturing firms and supply chains provide the productive base through which new technologies can be tested and adopted. Hydrogen creates potential routes into the decarbonisation of industrial heat, power and transport, whereas plastics and circularity research provide applications centred on material efficiency, recycling and resource recovery. Cells and quantum technologies may complement this configuration, but they should be treated as enabling capabilities rather than defining components of the cluster unless specific projects establish credible measurement, production-process or commercial applications.

The emergence of this cluster can therefore be understood as a process of industrial path renewal. Long-established universities, manufacturing firms and supply-chain relationships account for the persistence of the regional production base, while the expansion of transition-oriented themes since 2015 indicates that these inherited capabilities are increasingly being redirected towards decarbonisation and circular production. National place-based instruments and the emerging portfolio architecture provide an opportunity to coordinate these previously dispersed capabilities. At its current stage of development, however, the cluster requires stronger connections between research, demonstration and firm-level adoption. The next policy intervention should consequently prioritise shared demonstrator facilities, supplier upgrading and diffusion across the manufacturing base, with patient capital released against clearly defined scaling and adoption milestones rather than used to support another collection of disconnected research projects (Department for Business and Trade, 2025b; UKRI, 2026e; Tödtling & Trippl, 2005).

6.2 North West - Mission-Critical Science and AI-Enabled Materials Regional innovation cluster

The Northwest is best understood as a networked cluster of mission-oriented platforms connecting large-scale scientific infrastructure and critical industrial assets with materials, health and digital applications. Its distinctive configuration arises from the coexistence of exceptional nuclear specialisation with capabilities in cells, quantum technologies, astronomy, brain research and security. Together, these fields create opportunities for deep-tech instrumentation, advanced materials and applications that must operate within highly regulated or mission-critical environments.

The cluster is spatially and institutionally distributed. Liverpool contributes a substantial research and innovation base, Greater Manchester provides translational and portfolio-governance capacity, and the region’s large-science facilities and industrial assets offer environments in which new technologies can be tested and deployed. These nodes become

economically meaningful when they are connected through specific translational programmes. For example, the initial Local Innovation Partnerships Fund projects apply artificial intelligence to materials chemistry and biofilm research, while the Innovation Accelerator provides an additional mechanism for coordinating projects across institutional and sectoral boundaries (UKRI, 2026b; UKRI, 2026d). These initiatives provide early implementation evidence that the region is beginning to convert scientific diversity into a more integrated innovation portfolio.

The cluster has developed through two interacting processes. First, nationally significant scientific infrastructure and longstanding industrial specialisations have created a deep and relatively persistent capability base. Second, the expansion of city-region governance has strengthened the institutional capacity to connect these assets around shared missions. The cluster is therefore moving beyond capability accumulation towards coordinated implementation. Its principal weakness is spatial and institutional concentration: without deliberate intervention, benefits may remain confined to a limited number of major research institutions, large facilities and city-centre firms. Future funding should therefore attach supply-chain development, workforce participation and technology-adoption conditions to corridor-scale projects. Evaluation should correspondingly examine whether firms and communities beyond the strongest anchor institutions are participating in, and benefiting from, the cluster's development.

6.3 Scotland - Offshore Renewable Engineering and AI Systems Regional innovation cluster

Scotland's distinctive cluster connects offshore renewable energy and marine engineering with data-intensive design, monitoring, operation and maintenance. The joint prominence of wind turbines and vessels, together with a substantial climate-research base and the rapid growth of machine learning, indicates the emergence of an intelligent offshore-energy system rather than a generic green-innovation cluster.

This system contains three interconnected layers. The knowledge layer includes climate science, quantum technologies, cells and machine learning. The industrial layer is rooted in turbines, vessels, offshore engineering and associated supply chains. The application layer translates these capabilities into monitoring, grid integration, predictive maintenance and marine-environment decision-making. Machine learning is therefore not a separate regional specialisation but an enabling capability that can improve the design, efficiency and reliability of offshore-energy infrastructure. Similarly, climate research provides both the scientific rationale for the transition and the evidence needed to manage its environmental implications.

The formation of the cluster reflects the cumulative interaction of natural-resource opportunity, engineering history, university research and public strategy. Scotland's offshore conditions created the initial application context, while established engineering capabilities and research institutions supplied the knowledge and skills required to exploit it. More recently, Scotland's National Innovation Strategy and offshore wind policy have established a national governance

layer capable of coordinating assets that are distributed across multiple coastal and urban locations (Scottish Government, 2023; Scottish Government, 2020).

The cluster has consequently developed a relatively coherent relationship between research capability, industrial application and public policy. Its next maturity test, however, concerns regional value capture rather than the further accumulation of research activity. High research volume and extensive demonstrator assets will not, by themselves, establish a mature place-based ecosystem if deployment contracts, intellectual property, ownership, manufacturing and supply-chain benefits are captured elsewhere. Future policy should therefore assess not only technological outputs but also the extent to which innovation generates durable production, employment and investment within Scotland (Scottish Government, 2025).

6.4 Wales - Net-Zero Industrial Materials with a Bilingual Digital Layer

The Welsh cluster is centred on net-zero industrial materials, combining an established manufacturing and process-industry base with climate research, circular-material capabilities and low-carbon transition technologies. These elements support a coherent industrial proposition around metals, resource efficiency and the decarbonisation of production. Alongside this industrial core, Welsh-language technology constitutes a nationally distinctive but adjacent digital capability, with potential applications in public services, workplace interfaces and the inclusive adoption of new technologies.

The industrial component of the cluster is organised around a clear division of functions. Industrial sites and supply chains provide both the production base and the decarbonisation challenges to be addressed. Universities and research organisations contribute knowledge in climate science, materials and quantum-related fields, while circularity research connects waste streams, material recovery and resource-efficient production. The Welsh innovation strategy supplies a mission-oriented coordination layer through which these research, industrial and public-sector capabilities can be aligned (Welsh Government, 2023a; Welsh Government, 2023b).

The cluster has emerged through the renewal of an inherited industrial path. Wales's industrial history explains the persistence of its production and process-engineering capabilities, whereas the diversification of research themes since 2015 reflects growing decarbonisation pressures and the broadening of the national innovation agenda. The bilingual digital layer arises from a different but potentially complementary institutional context. It should therefore be incorporated into the industrial proposition only where Welsh-language technologies address identifiable user, workforce or public-service requirements. Treating it automatically as part of the same cluster would risk substituting narrative coherence for demonstrated functional connection.

The maturity assessment should consequently remain differentiated. The industrial component already displays meaningful connections among production assets, research capabilities and

policy priorities. By contrast, the bilingual digital layer requires targeted experimentation to establish its additionality, adoption pathways and connection to industrial or public-service demand. Maintaining this distinction allows Wales to build a place-based strategy around complementary strengths without forcing every distinctive regional capability into a single, artificially unified cluster narrative.

6.5 West Midlands - Manufacturing-Health-AI Convergence Regional innovation cluster

The West Midlands cluster represents the reorientation of an established engineering and production system through data-intensive technologies, health applications and service innovation. Manufacturing remains the region's principal applied capability, while the combination of cells, care and machine learning creates an opportunity for cross-sectoral convergence. The most analytically defensible proposition is therefore intelligent production for health and mobility rather than a generic artificial-intelligence cluster.

The cluster's internal structure reflects the complementary functions of different regional actors. Coventry and Warwick connect advanced manufacturing, engineering research and industrial firms. Birmingham contributes clinical research, a large health-service system and urban demand capable of supporting experimentation and adoption. Firms provide the production capabilities and commercial routes through which innovations can be scaled, while the Innovation Accelerator offers a portfolio mechanism to connect research, demonstration, and translation. Artificial intelligence is valuable within this configuration because it can link production data, clinical requirements and mobility systems; it does not constitute a self-sufficient regional cluster in the absence of these application contexts.

This configuration extends the direction established by the Local Industrial Strategy, which sought to connect the region's manufacturing heritage, research assets and devolved productivity agenda (West Midlands Combined Authority & Department for Business, Energy and Industrial Strategy, 2019; West Midlands Combined Authority, n.d.). Its development is supported by both supply- and demand-side conditions: industrial depth and university capabilities create the technological supply, while the regional health system, mobility challenges and devolution arrangements create demand and coordination capacity.

The principal barrier to further maturity is no longer the identification of promising technologies but their repeated adoption across organisations and supply chains. One-off demonstrators will have limited regional impact unless they generate procurement, production and diffusion beyond the initial project participants. Funding should therefore combine challenge-led procurement, shared demonstrator facilities and scale-up finance within a single intervention sequence. Local supplier participation and repeated adoption should be treated as explicit performance dimensions rather than assumed consequences of research and development expenditure.

6.6 Northern Ireland - Secure Connected Systems and Green-Mobility Renewal

Northern Ireland's emerging cluster is distinguished by the strong relative presence of security research alongside wireless and laser capabilities. At the same time, rising hydrogen activity and declining shares in some established digital and creative themes suggest that the regional innovation base is undergoing renewal rather than simply extending its previous trajectory. The resulting proposition is a forming cluster of secure connected systems applied to health, public services, mobility and green transport.

The cluster combines a set of potentially complementary technical and application capabilities. Security, communications and photonics form the technological base, while health, public services and transport provide mission-oriented demand. Hydrogen offers a possible route into green mobility. Still, it should be incorporated into the cluster only when production or distribution infrastructure, end users and demonstrator projects are jointly organised. Geographically, Belfast provides the largest research and business node, while the Derry/Londonderry corridor broadens the cluster's functional reach and reduces the risk that innovation activity is treated as a single-city phenomenon.

This emerging configuration is consistent with Northern Ireland's evolving policy architecture. The 10X vision connects innovation with inclusion and sustainability, while the technologies-and-clusters framework identifies enabling technologies and priority sectors through which these objectives may be pursued (Department for the Economy, 2021; Department for the Economy, 2023). These strategies provide policy recognition and a basis for coordination, but they do not in themselves demonstrate that a stable innovation ecosystem has formed.

The cluster remains distinctive but volatile. It possesses identifiable capabilities and plausible application domains, yet the connections among firms, research organisations, public demand and investment are not sufficiently stable to support a claim of mature integration. Its funding model should therefore reflect this developmental stage. Initial support should establish institutional connections and credible use cases; subsequent investment should develop shared demonstrators and firm-level absorptive capacity; and stronger private-match, adoption and scale requirements should be introduced only after these foundations have been demonstrated. This sequential approach is consistent with the need to differentiate regional innovation policy according to the institutional and connectivity constraints of the place concerned (Tödtling & Trippl, 2005).

7. Concluding remarks

This report set out to examine whether UKRI-funded research projects exhibit meaningful thematic differentiation across UK regions, and how that differentiation has evolved over time. The evidence points to a consistent and somewhat paradoxical picture. At the national level, a small number of broad themes — society, policy, culture and gender; climate, plants, crops and species; cells, cancer, proteins and infection; and quantum technologies, devices and

batteries — dominate the funding portfolio in every region, and this core structure has remained remarkably stable between 2006 and 2024. Regional differentiation, where it exists, is not found in the largest topics driving each region's portfolio, but in secondary themes that recur consistently within specific university cities and industrial ecosystems. These secondary specialisations are sufficiently coherent to support the identification of six distinctive emerging innovations. Importantly, these clusters appear to represent processes of related diversification around inherited industrial and research strengths, rather than the emergence of entirely new regional trajectories, and most remain at an early or intermediate stage of maturity rather than fully realised, self-sustaining ecosystems.

7.1 Limitations and future research

As an initial step to provide evidence for the effectiveness of place-based innovation policies, this study has limitations. Addressing some of these aspects, however, may open interesting research avenues and reveal additional insights. First, our analysis of funded project abstracts may not fully represent innovation capabilities across regions. As the next step, combining the abstract with more 'output' data would be ideal, for example, including patent analysis for technologically orientated projects and trademarks for creativity-related projects. In addition, it would be interesting to include the actual amount the project received, so we can add some sort of 'weighting' to the topics analysed. Currently, the topics are discussed based on counts, without considering the quantity of financial support they received.

Second, the analysis revealed many connections between cities and regions, but we could only partially capture them through interpretation. A more systematic exploration regarding how different institutions work together may reveal further insights regarding the existence of innovation clusters. For example, institutions may collaborate outside of their regions based on reputation and other capability considerations. In this vein, innovation clusters may take an intangible network form rather than being geographically based. We cannot capture this network effect, but we thought it would lead to interesting discoveries.

7.2 Policy implications

First, Local Innovation Partnerships Fund (LIPF) awards should be directed towards the secondary specialisations identified here, not the shared national themes. The £500 million LIPF (2026–2031) is explicitly structured around locally defined "high-potential" clusters, with earmarked awards already confirmed for Greater Manchester, the West Midlands, South and West Yorkshire, Liverpool City Region, the North East, London, Glasgow, Cardiff and the Belfast–Derry/Londonderry corridor, and a competed strand offering up to £20 million per place for clusters not covered by an earmarked award. Because this report shows that regional distinctiveness sits in secondary themes rather than the largest topic in each region, LIPF partnerships should use portfolio-selection criteria that explicitly privilege these secondary specialisations, for example, hydrogen and circular manufacturing in the East Midlands, or

nuclear-and-astronomy-adjacent instrumentation in the Northwest, rather than defaulting to the nationally dominant themes that already attract the bulk of mainstream UKRI funding.

Second, further aligning regional secondary specialisations explicitly with the Modern Industrial Strategy's IS-8 sector plans would create stronger cluster effects in both innovation and productivity gains. The 2025 Modern Industrial Strategy identifies eight priority sectors, namely advanced manufacturing, clean energy, creative industries, defence, digital and technologies, financial services, life sciences, and professional and business services, each with a detailed sector plan published in June 2025. Several of the regional specialisations identified in this report map directly onto these sectors: Scotland's offshore renewable engineering cluster onto the Clean Energy Industries sector plan's commitments to offshore wind and hydrogen hubs; the South West and South East's AI-enabled clusters onto the Digital and Technologies sector plan's focus on AI Growth Zones; and the East Midlands and Wales's manufacturing clusters onto the Advanced Manufacturing sector plan's support for aerospace, automotive, batteries and advanced materials. Regional funding bodies should explicitly reference the relevant sector plan when developing LIPF bids to strengthen the case for national co-investment alongside UKRI funding.

Third, differentiate the funding model by cluster maturity, following the staged model UKRI and the British Business Bank already piloted. The clusters identified in this report sit at markedly different stages of development, and current UK funding architecture already recognises, at least in principle, that a single instrument cannot serve all of them equally well. The financial support should be tailored more explicitly to the clusters identified here. More established clusters, for example, Scotland's offshore renewable engineering cluster and the South West's AI-enabled manufacturing clusters, are already showing readiness for scale-up finance and stronger private-match conditions. By contrast, emerging clusters, such as Northern Ireland's secure connected systems cluster, need earlier-stage support focused on building institutional connections and credible use cases before adoption-based or match-funding conditions are introduced.

Taken together, these implications suggest that central policy priorities can move beyond identifying where research strength exists to building the coordination, infrastructure, and adoption mechanisms needed to convert that strength into durable, regionally anchored economic and social benefit.

REFERENCES

- Adner, R. (2017). Regional innovation cluster as structure: An actionable construct for strategy. *Journal of Management*, 43(1), 39–58. <https://doi.org/10.1177/0149206316678451>
- Arts and Humanities Research Council. (2022). *Strategic delivery plan 2022 to 2025*. <https://www.ukri.org/publications/ahrc-strategic-delivery-plan/ahrc-strategic-delivery-plan-2022-to-2025/>
- Asheim, B. T., Smith, H. L., & Oughton, C. (2011). Regional innovation systems: Theory, empirics and policy. *Regional Studies*, 45(7), 875–891. <https://doi.org/10.1080/00343404.2011.596701>
- Balland, P.-A., Boschma, R., Crespo, J., & Rigby, D. L. (2019). Smart specialization policy in the European Union: Relatedness, knowledge complexity and regional diversification. *Regional Studies*, 53(9), 1252–1268. <https://doi.org/10.1080/00343404.2018.1437900>
- Barca, F., McCann, P., & Rodríguez-Pose, A. (2012). The case for regional development intervention: Place-based versus place-neutral approaches. *Journal of Regional Science*, 52(1), 134–152. <https://doi.org/10.1111/j.1467-9787.2011.00756.x>
- Biotechnology and Biological Sciences Research Council. (2022). *Strategic delivery plan 2022 to 2025*. <https://www.ukri.org/publications/bbsrc-strategic-delivery-plan/bbsrc-strategic-delivery-plan-2022-to-2025/>
- Blei, D. M., Ng, A. Y., & Jordan, M. I. 2003. Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3: 993–1022.
- Boschma, R. A. (2005). Proximity and innovation: A critical assessment. *Regional Studies*, 39(1), 61–74. <https://doi.org/10.1080/0034340052000320887>
- Campello, R. J. G. B., Moulavi, D., & Sander, J. 2013. Density-based clustering based on hierarchical density estimates. In *Advances in Knowledge Discovery and Data Mining*: 160–172. Berlin: Springer.
- Campello, R. J. G. B., Moulavi, D., Zimek, A., & Sander, J. 2015. Hierarchical density estimates for data clustering, visualization, and outlier detection. *ACM Transactions on Knowledge Discovery from Data*, 10(1): Article 5.
- Cooke, P., Uranga, M. G., & Etxebarria, G. (1997). Regional innovation systems: Institutional and organisational dimensions. *Research Policy*, 26(4–5), 475–491. [https://doi.org/10.1016/S0048-7333\(97\)00025-5](https://doi.org/10.1016/S0048-7333(97)00025-5)
- Department for Business and Trade. (2025a, June 23). *Industrial strategy sector plans*. <https://www.gov.uk/government/publications/industrial-strategy-sector-plans/sector-plans>
- Department for Business and Trade. (2025b, June 23). *The UK's modern industrial strategy*. <https://www.gov.uk/government/publications/industrial-strategy>

Department for Business and Trade. (2025c). *The UK's modern industrial strategy: Yorkshire and the Humber*. https://www.business.gov.uk/documents/1419/IS_Yorkshire_and_the_Humber.pdf

Department for Business, Energy and Industrial Strategy. (2017, November 27). *Industrial strategy: Building a Britain fit for the future*. <https://www.gov.uk/government/publications/industrial-strategy-building-a-britain-fit-for-the-future>

Department for Business, Energy and Industrial Strategy. (2020, July 1). *UK research and development roadmap*. <https://www.gov.uk/government/publications/uk-research-and-development-roadmap/uk-research-and-development-roadmap>

Department for Business, Energy and Industrial Strategy. (2021, July 22). *UK innovation strategy: Leading the future by creating it*. <https://www.gov.uk/government/publications/uk-innovation-strategy-leading-the-future-by-creating-it/uk-innovation-strategy-leading-the-future-by-creating-it-accessible-webpage>

Department for Business, Enterprise and Regulatory Reform. (2008, March 13). *Innovation nation*. <https://www.gov.uk/government/publications/innovation-nation>

Department for Business, Innovation and Skills. (2015, November 19). *Nurse review of the research councils: Recommendations*. <https://www.gov.uk/government/publications/nurse-review-of-research-councils-recommendations>

Department for Levelling Up, Housing and Communities. (2022, February 2). *Levelling up the United Kingdom*. <https://www.gov.uk/government/publications/levelling-up-the-united-kingdom>

Department for Levelling Up, Housing and Communities & HM Treasury. (2023, March 15). *Investment Zones policy prospectus*. <https://www.gov.uk/government/publications/investment-zones/investment-zones-policy-prospectus>

Department for Science, Innovation and Technology. (2023, March 6). *The UK science and technology framework*. <https://www.gov.uk/government/publications/uk-science-and-technology-framework/the-uk-science-and-technology-framework>

Department for Science, Innovation and Technology. (2025a, October 1). *Innovation Clusters Map: Summary and methods*. <https://www.gov.uk/government/publications/innovation-clusters-map-summary-and-methods/innovation-clusters-map-summary-and-methods>

Department for Science, Innovation and Technology. (2025b, April 28). *Science and technology framework*. <https://www.gov.uk/government/publications/science-and-technology-framework/science-and-technology-framework>

Department for the Economy. (2021, May). *10X economy: An economic vision for a decade of innovation*. <https://www.economy-ni.gov.uk/articles/10x-economy-economic-vision>

- Department for the Economy. (2023, September 29). *Approach to 10X technologies and clusters*. <https://www.economy-ni.gov.uk/consultations/approach-10x-technologies-and-clusters>
- Economic and Social Research Council. (2022a, October 6). *Developing Local Policy Innovation Partnerships*. <https://www.ukri.org/opportunity/developing-local-policy-innovation-partnerships/>
- Economic and Social Research Council. (2022b). *Strategic delivery plan 2022 to 2025*. <https://www.ukri.org/publications/esrc-strategic-delivery-plan/esrc-strategic-delivery-plan-2022-to-2025/>
- Engineering and Physical Sciences Research Council. (2022). *Strategic delivery plan 2022 to 2025*. <https://www.ukri.org/publications/epsrc-strategic-delivery-plan/epsrc-strategic-delivery-plan-2022-to-2025/>
- Engineering and Physical Sciences Research Council. (2025, December 15). *Place-Based Impact Acceleration Account early-stage review: Evaluation report*. <https://www.ukri.org/publications/place-based-iaa-early-stage-review-evaluation-report/>
- Flanagan, K., Uyarra, E., & Laranja, M. (2011). Reconceptualising the ‘policy mix’ for innovation. *Research Policy*, 40(5), 702–713. <https://doi.org/10.1016/j.respol.2011.02.005>
- Grimmer, J., & Stewart, B. M. 2013. Text as data: The promise and pitfalls of automatic content analysis methods for political texts. *Political Analysis*, 21(3): 267–297.
- Grootendorst, M. 2022. BERTopic: Neural topic modeling with a class-based TF-IDF procedure. arXiv preprint arXiv:2203.05794.
- Hannigan, T. R., Haans, R. F. J., Vakili, K., Tchalian, H., Glaser, V. L., Wang, M. S., Kaplan, S., & Jennings, P. D. 2019. Topic modeling in management research: Rendering new theory from textual data. *Academy of Management Annals*, 13(2): 586–632.
- HM Treasury & Department for Business, Innovation and Skills. (2014, December). *Our plan for growth: Science and innovation*. https://assets.publishing.service.gov.uk/media/5a7d0003e5274a33be648537/PU1719_HMT_Science.pdf
- Innovate UK. (2022, September 30). *Strategic delivery plan 2022 to 2025*. <https://www.ukri.org/publications/innovate-uk-strategic-delivery-plan/innovate-uk-strategic-delivery-plan-2022-to-2025/>
- McInnes, L., Healy, J., & Melville, J. 2018. UMAP: Uniform Manifold Approximation and Projection for dimension reduction. arXiv preprint arXiv:1802.03426.
- Medical Research Council. (2022). *Strategic delivery plan 2022 to 2025*. <https://www.ukri.org/publications/mrc-strategic-delivery-plan/mrc-strategic-delivery-plan-2022-to-2025/>

- Ministry of Housing, Communities and Local Government. (2024, December 16). *English devolution white paper: Power and partnership—Foundations for growth*. <https://www.gov.uk/government/publications/english-devolution-white-paper-power-and-partnership-foundations-for-growth/english-devolution-white-paper>
- Ministry of Housing, Communities and Local Government. (2025, June 23). *Industrial Strategy Zones Action Plan*. <https://www.gov.uk/government/publications/industrial-strategy-zones-action-plan/industrial-strategy-zones-action-plan>
- Natural Environment Research Council. (2022). *Strategic delivery plan 2022 to 2025*. <https://www.ukri.org/publications/nerc-strategic-delivery-plan/nerc-strategic-delivery-plan-2022-to-2025/>
- Reimers, N., & Gurevych, I. 2019. Sentence-BERT: Sentence embeddings using Siamese BERT-networks. Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing: 3982–3992.
- Research England. (2021, September 1). *Strength in Places Fund evaluation framework*. <https://www.ukri.org/publications/strength-in-places-fund-evaluation-framework/>
- Research England. (2022a). *Strategic delivery plan 2022 to 2025*. <https://www.ukri.org/publications/research-england-strategic-delivery-plan/research-england-strategic-delivery-plan-sdp-2022-to-2025/>
- Research England. (2022b, October 1). *Strength in Places Fund evaluation baseline report*. <https://www.ukri.org/publications/strength-in-places-fund-evaluation-baseline-report/>
- Research England. (2023, October 25). *Regional Innovation Fund 2023 to 2024*. <https://www.ukri.org/publications/regional-innovation-fund-2023-to-2024/>
- Research England. (2025a, February 7). *Final evaluation of the Connecting Capability Fund programme*. <https://www.ukri.org/publications/final-evaluation-of-the-connecting-capability-fund-ccf-programme/>
- Research England. (2025b, October 29). *Higher Education Innovation Funding policies and priorities 2025 to 2031*. <https://www.ukri.org/publications/research-england-heif-policies-and-priorities/higher-education-innovation-funding-policies-and-priorities-2025-to-2031/>
- Research England. (2025c, March 11). *Strength in Places Fund interim process evaluation reports*. <https://www.ukri.org/publications/strength-in-places-fund-interim-process-evaluation-reports/>
- Research England. (2026a, May 7). *Connecting Capability Fund*. <https://www.ukri.org/what-we-do/browse-our-areas-of-investment-and-support/connecting-capability-fund/>
- Research England. (2026b, May 29). *Higher Education Innovation Fund*. <https://www.ukri.org/what-we-do/browse-our-areas-of-investment-and-support/higher-education-innovation-fund/>

- Science and Technology Facilities Council. (2022). *Strategic delivery plan 2022 to 2025*. <https://www.ukri.org/publications/stfc-strategic-delivery-plan/stfc-strategic-delivery-plan-2022-to-2025/>
- Scottish Government. (2020, October 28). *Offshore wind policy statement*. <https://www.gov.scot/publications/offshore-wind-policy-statement/pages/8/>
- Scottish Government. (2023, June 9). *Scotland's national innovation strategy 2023 to 2033*. <https://www.gov.scot/publications/scotlands-national-innovation-strategy/>
- Scottish Government. (2025, September 25). *National innovation strategy scorecard: 2025 update*. <https://www.gov.scot/publications/national-innovation-strategy-scorecard-2025-update/>
- Tödting, F., & Trippel, M. (2005). One size fits all? Towards a differentiated regional innovation policy approach. *Research Policy*, 34(8), 1203–1219. <https://doi.org/10.1016/j.respol.2005.01.018>
- UK Research and Innovation. (2021, October 21). *Strength in Places Fund: Funded projects*. <https://www.ukri.org/publications/strength-in-places-fund-funded-projects/>
- UK Research and Innovation. (2022a, August 22). *UKRI corporate plan 2022 to 2025*. <https://www.ukri.org/publications/ukri-corporate-plan/ukri-corporate-plan-2022-to-2025/>
- UK Research and Innovation. (2022b, March 17). *UKRI strategy 2022 to 2027*. <https://www.ukri.org/publications/ukri-strategy-2022-to-2027/ukri-strategy-2022-to-2027/>
- UK Research and Innovation. (2023a, January 5). *Strength in Places Fund*. <https://www.ukri.org/what-we-do/browse-our-areas-of-investment-and-support/strength-in-places-fund/>
- UK Research and Innovation. (2023b, October 6). *UKRI corporate plan update 2023 to 2024*. <https://www.ukri.org/publications/ukri-corporate-plan-update-2023-to-2024/uk-research-and-innovation-corporate-plan-2023-to-2024-update/>
- UK Research and Innovation. (2024a, November 19). *Research and innovation clusters to boost UK regional growth*. <https://www.ukri.org/news/research-and-innovation-clusters-to-boost-uk-regional-growth/>
- UK Research and Innovation. (2024b, August 29). *UKRI evaluation strategy*. <https://www.ukri.org/publications/uk-research-and-innovation-evaluation-strategy/ukri-evaluation-strategy/>
- UK Research and Innovation. (2025a, March 10). *Geographical distribution of UKRI funding: Financial years 2022 to 2023 and 2023 to 2024*. <https://www.ukri.org/publications/2022-to-2023-and-2023-to-2024-geographical-distribution-of-funding/geographical-distribution-of-ukri-funding-financial-years-2022-to-2023-and-2023-to-2024/>
- UK Research and Innovation. (2025b, October 6). *Local Innovation Partnerships Fund: Competed strand*. <https://www.ukri.org/opportunity/local-innovation-partnerships-fund/>

- UK Research and Innovation. (2025c, November 25). *UKRI corporate plan update 2025 to 2027*. <https://www.ukri.org/publications/ukri-corporate-plan-update-2025-to-2027/ukri-corporate-plan-update-2025-to-2027/>
- UK Research and Innovation. (2026a, August 12). *Annual report and accounts 2025 to 2026*. <https://www.ukri.org/publications/annual-report-and-accounts-2025-to-2026/ukri-annual-report-and-accounts-2025-to-2026/>
- UK Research and Innovation. (2026b, June 1). *First projects funded by Local Innovation Partnership Fund*. <https://www.ukri.org/news/first-projects-funded-by-local-innovation-partnerships-fund/>
- UK Research and Innovation. (2026c, March 25). *Innovation Accelerator pilot set to create lasting legacy*. <https://www.ukri.org/news/innovation-accelerator-pilot-set-to-create-lasting-legacy/>
- UK Research and Innovation. (2026d, January 23). *Innovation Accelerator programme*. <https://www.ukri.org/what-we-do/browse-our-areas-of-investment-and-support/innovation-accelerator-programme/>
- UK Research and Innovation. (2026e, July 31). *Local Innovation Partnerships Fund*. <https://www.ukri.org/what-we-do/browse-our-areas-of-investment-and-support/local-innovation-partnerships-fund/>
- UK Research and Innovation. (2026f, July 13). *Our objectives and key results 2026 to 2027*. <https://www.ukri.org/publications/uk-research-and-innovation-delivery-plan-2026-to-2027/our-objectives-and-key-results-2026-to-2027/>
- UK Research and Innovation. (2026g, April 2). *Seven new UK areas confirmed for £500m innovation fund support*. <https://www.ukri.org/news/seven-new-uk-areas-confirmed-for-500m-innovation-fund-support/>
- UK Research and Innovation. (2026h, July 13). *UKRI delivery plan 2026 to 2027*. <https://www.ukri.org/publications/uk-research-and-innovation-delivery-plan-2026-to-2027/ukri-delivery-plan-2026-to-2027/>
- UK Research and Innovation. (2026i, July 16). *UKRI investment and outputs: April 2015 to March 2026*. <https://www.ukri.org/publications/ukri-investments-and-outputs/statistical-publication-ukri-investment-and-outputs-april-2015-to-march-2026/>
- UK Research and Innovation. (2026j, July 13). *UKRI strategy 2026 to 2031*. <https://www.ukri.org/publications/uk-research-and-innovation-strategy-2026-to-2031/ukri-strategy-2026-to-2031/>
- Welsh Government. (2023a, February 27). *Wales innovates: Creating a stronger, fairer, greener Wales*. <https://www.gov.wales/wales-innovates-creating-stronger-fairer-greener-wales-html>
- Welsh Government. (2023b, December). *Wales innovates: Delivery plan*. <https://www.gov.wales/innovation-strategy-wales-delivery-plan>

West Midlands Combined Authority. (n.d.). *West Midlands Innovation Accelerator*.
<https://www.wmca.org.uk/what-we-do/economy-and-innovation/west-midlands-innovation/west-midlands-innovation-accelerator/>

West Midlands Combined Authority & Department for Business, Energy and Industrial Strategy. (2019, May 16). *West Midlands local industrial strategy*.
<https://www.gov.uk/government/publications/west-midlands-local-industrial-strategy/west-midlands-local-industrial-strategy>

APPENDIX

Appendix A. Full topic descriptions

Table A.1 Full list of topics with five most important terms

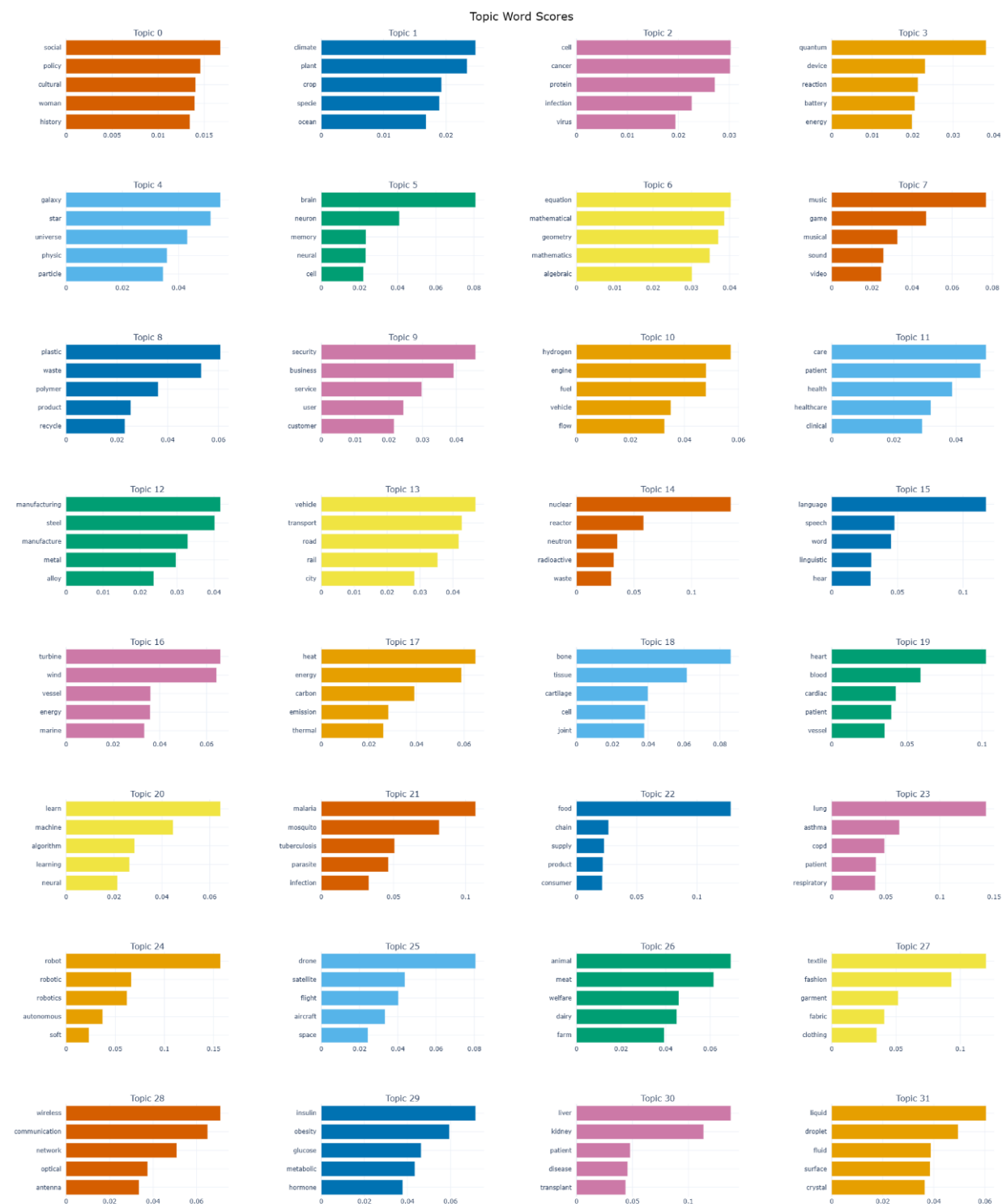
ID	Topic names	Documents	Rate	Most important terms
0	social_policy_cultural_woman	18921	15	social, policy, cultural, woman, history
1	climate_plant_crop_specie	12571	9.96	climate, plant, crop, specie, ocean
2	cell_cancer_protein_infection	11731	9.3	cell, cancer, protein, infection, virus
3	quantum_device_reaction_battery	10746	8.52	quantum, device, reaction, battery, energy
4	galaxy_star_universe_physics	4490	3.56	galaxy, star, universe, physic, particle
5	brain_neuron_memory_neural	4175	3.31	brain, neuron, memory, neural, cell
6	equation_mathematical_geometry_mathematics	3165	2.51	equation, mathematical, geometry, mathematics, algebraic
7	music_game_musical_sound	2892	2.29	music, game, musical, sound, video
8	plastic_waste_polymer_product	2889	2.29	plastic, waste, polymer, product, recycle
9	security_business_service_user	4433	3.51	security, business, service, user, customer
10	hydrogen_engine_fuel_vehicle	2884	2.29	hydrogen, engine, fuel, vehicle, flow
11	care_patient_health_healthcare	3187	2.53	care, patient, health, healthcare, clinical
12	manufacturing_steel_manufacture_metal	2609	2.07	manufacturing, steel, manufacture, metal, alloy
13	vehicle_transport_road_rail	2025	1.6	vehicle, transport, road, rail, city
14	nuclear_reactor_neutron_radioactive	1411	1.12	nuclear, reactor, neutron, radioactive, waste
15	language_speech_word_linguistic	1691	1.34	language, speech, word, linguistic, hear
16	turbine_wind_vessel_energy	1607	1.27	turbine, wind, vessel, energy, marine
17	heat_energy_carbon_emission	2436	1.93	heat, energy, carbon, emission, thermal
18	bone_tissue_cartilage_cell	1200	0.95	bone, tissue, cartilage, cell, joint
19	heart_blood_cardiac_patient	1178	0.93	heart, blood, cardiac, patient, vessel

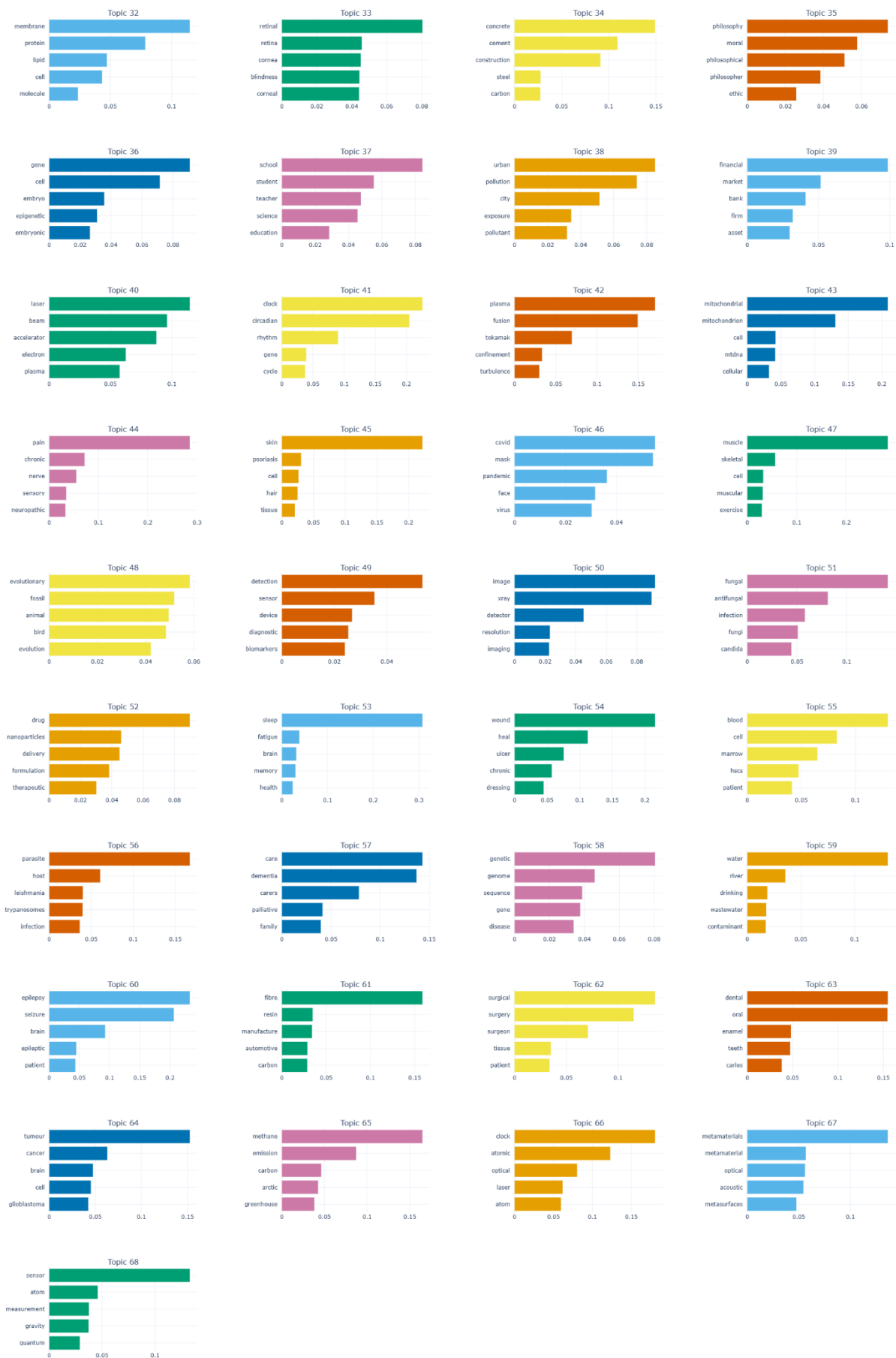
ID	Topic names	Documents	Rate	Most important terms
20	learn_machine_algorithm_learning	2162	1.71	learn、machine、algorithm、learning、neural
21	malaria_mosquito_tuberculosis_parasite	738	0.58	malaria、mosquito、tuberculosis、parasite、infection
22	food_chain_supply_product	1741	1.38	food、chain、supply、product、consumer
23	lung_asthma_copd_patient	746	0.59	lung、asthma、copd、patient、respiratory
24	robot_robotic_robotics_autonomous	811	0.64	robot、robotic、robotics、autonomous、soft
25	drone_satellite_flight_aircraft	789	0.63	drone、satellite、flight、aircraft、space
26	animal_meat_welfare_dairy	849	0.67	animal、meat、welfare、dairy、farm
27	textile_fashion_garment_fabric	723	0.57	textile、fashion、garment、fabric、clothing
28	wireless_communication_network_optical	989	0.78	wireless、communication、network、optical、antenna
29	insulin_obesity_glucose_metabolic	701	0.56	insulin、obesity、glucose、metabolic、hormone
30	liver_kidney_patient_disease	552	0.44	liver、kidney、patient、disease、transplant
31	liquid_droplet_fluid_surface	979	0.78	liquid、droplet、fluid、surface、crystal
32	membrane_protein_lipid_cell	998	0.79	membrane、protein、lipid、cell、molecule
33	retinal_retina_cornea_blindness	386	0.31	retinal、retina、cornea、blindness、corneal
34	concrete_cement_construction_steel	472	0.37	concrete、cement、construction、steel、carbon
35	philosophy_moral_philosophical_philosopher	623	0.49	philosophy、moral、philosophical、philosopher、ethic
36	gene_cell_embryo_epigenetic	1043	0.83	gene、cell、embryo、epigenetic、embryonic
37	school_student_teacher_science	1080	0.86	school、student、teacher、science、education
38	urban_pollution_city_exposure	884	0.7	urban、pollution、city、exposure、pollutant
39	financial_market_bank_firm	815	0.65	financial、market、bank、firm、asset
40	laser_beam_accelerator_electron	467	0.37	laser、beam、accelerator、electron、plasma

ID	Topic names	Documents	Rate	Most important terms
41	clock_circadian_rhythm_gene	242	0.19	clock, circadian, rhythm, gene, cycle
42	plasma_fusion_tokamak_confinement	314	0.25	plasma, fusion, tokamak, confinement, turbulence
43	mitochondrial_mitochondrion_cell_mtdna	250	0.2	mitochondrial, mitochondrion, cell, mtdna, cellular
44	pain_chronic_nerve_sensory	355	0.28	pain, chronic, nerve, sensory, neuropathic
45	skin_psoriasis_cell_hair	420	0.33	skin, psoriasis, cell, hair, tissue
46	covid_mask_pandemic_face	413	0.33	covid, mask, pandemic, face, virus
47	muscle_skeletal_cell_muscular	311	0.25	muscle, skeletal, cell, muscular, exercise
48	evolutionary_fossil_animal_bird	432	0.34	evolutionary, fossil, animal, bird, evolution
49	detection_sensor_device_diagnostic	902	0.71	detection, sensor, device, diagnostic, biomarkers
50	image_xray_detector_resolution	719	0.57	image, xray, detector, resolution, imaging
51	fungus_antifungal_infection_fungi	267	0.21	fungus, antifungal, infection, fungi, candida
52	drug_nanoparticles_delivery_formulation	929	0.74	drug, nanoparticles, delivery, formulation, therapeutic
53	sleep_fatigue_brain_memory	204	0.16	sleep, fatigue, brain, memory, health
54	wound_heal_ulcer_chronic	229	0.18	wound, heal, ulcer, chronic, dressing
55	blood_cell_marrow_hscs	334	0.26	blood, cell, marrow, hscs, patient
56	parasite_host_leishmania_trypanosomes	303	0.24	parasite, host, leishmania, trypanosomes, infection
57	care_dementia_carers_palliative	329	0.26	care, dementia, carers, palliative, family
58	genetic_genome_sequence_gene	858	0.68	genetic, genome, sequence, gene, disease
59	water_river_drinking_wastewater	1074	0.85	water, river, drinking, wastewater, contaminant
60	epilepsy_seizure_brain_epileptic	150	0.12	epilepsy, seizure, brain, epileptic, patient
61	fibre_resin_manufacture_automotive	414	0.33	fibre, resin, manufacture, automotive, carbon
62	surgical_surgery_surgeon_tissue	248	0.2	surgical, surgery, surgeon, tissue, patient

ID	Topic names	Documents	Rate	Most important terms
63	dental_oral_enamel_teeth	174	0.14	dental, oral, enamel, teeth, caries
64	tumour_cancer_brain_cell	285	0.23	tumour, cancer, brain, cell, glioblastoma
65	methane_emission_carbon_arctic	342	0.27	methane, emission, carbon, arctic, greenhouse
66	clock_atomic_optical_laser	177	0.14	clock, atomic, optical, laser, atom
67	metamaterials_metamaterial_optical_acoustic	217	0.17	metamaterials, metamaterial, optical, acoustic, metasurfaces
68	sensor_atom_measurement_gravity	500	0.4	sensor, atom, measurement, gravity, quantum

Figure A.1 Topic keyword scores





Appendix B. Top 10 Topics per Region

Table B.1 East Midlands

Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	1,042	17.5%
2	Climate / plants / crops / species	623	10.5%
3	Cells / cancer / proteins / infection	593	10.0%
4	Quantum / devices / reactions / batteries	525	8.8%
5	Brain / neurons / memory / neuroscience	209	3.5%
6	Galaxies / stars / universe / physics	205	3.4%
7	Hydrogen / engines / fuels / vehicles	190	3.2%
8	Plastics / waste / polymers / products	169	2.8%
9	Equations / mathematics / geometry	144	2.4%
10	Manufacturing / steel / metals	142	2.4%

Table B.2 East of England

Rank	Topic	Topic records	Share within region
1	Climate / plants / crops / species	1,624	18.7%
2	Society / policy / culture / women	1,105	12.7%
3	Cells / cancer / proteins / infection	1,009	11.6%
4	Quantum / devices / reactions / batteries	760	8.8%
5	Galaxies / stars / universe / physics	418	4.8%
6	Brain / neurons / memory / neuroscience	359	4.1%
7	Equations / mathematics / geometry	229	2.6%
8	Genes / cells / embryos / epigenetics	195	2.2%
9	Hydrogen / engines / fuels / vehicles	172	2.0%
10	Machine learning / algorithms	150	1.7%

Table B.3 London

Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	3,779	20.0%
2	Cells / cancer / proteins / infection	2,298	12.2%
3	Quantum / devices / reactions / batteries	1,239	6.6%
4	Brain / neurons / memory / neuroscience	1,186	6.3%
5	Climate / plants / crops / species	1,082	5.7%
6	Galaxies / stars / universe / physics	686	3.6%
7	Equations / mathematics / geometry	567	3.0%
8	Care / patients / healthcare***	475	2.5%
9	Machine learning / algorithms***	407	2.2%
10	Security / business / services / users***	369	2.0%

Table B.4 North East

Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	959	21.6%
2	Climate / plants / crops / species	560	12.6%
3	Quantum / devices / reactions / batteries	457	10.3%
4	Cells / cancer / proteins / infection	359	8.1%
5	Galaxies / stars / universe / physics	256	5.8%
6	Equations / mathematics / geometry	132	3.0%
7	Brain / neurons / memory / neuroscience	102	2.3%
8	Security / business / services / users***	99	2.2%
9	Plastics / waste / polymers / products***	88	2.0%
10	Hydrogen / engines / fuels / vehicles***	75	1.7%

Table B.5 North West

Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	1,348	15.1%
2	Cells / cancer / proteins / infection	882	9.9%
3	Quantum / devices / reactions / batteries	839	9.4%
4	Climate / plants / crops / species	699	7.8%
5	Galaxies / stars / universe / physics	584	6.5%
6	Nuclear energy / reactors / neutrons / radioactivity***	363	4.1%
7	Brain / neurons / memory / neuroscience	268	3.0%
8	Equations / mathematics / geometry	218	2.4%
9	Care / patients / healthcare***	181	2.0%
10	Security / business / services / users***	179	2.0%

Table B.6 Northern Ireland

Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	358	24.9%
2	Cells / cancer / proteins / infection	144	10.0%
3	Quantum / devices / reactions / batteries	117	8.1%
4	Climate / plants / crops / species	88	6.1%
5	Galaxies / stars / universe / physics	75	5.2%
6	Security / business / services / users***	53	3.7%
7	Lasers / beams / accelerators / electrons***	39	2.7%
8	Music / games / sound***	36	2.5%
9	Hydrogen / engines / fuels / vehicles***	31	2.2%
10	Wireless communications / networks / optics***	28	1.9%

Table B.7 Scotland

Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	1,683	13.8%
2	Climate / plants / crops / species	1,506	12.3%
3	Cells / cancer / proteins / infection	1,380	11.3%
4	Quantum / devices / reactions / batteries	1,215	10.0%
5	Galaxies / stars / universe / physics	476	3.9%
6	Brain / neurons / memory / neuroscience	462	3.8%
7	Equations / mathematics / geometry	397	3.3%
8	Turbines / wind energy / vessels / energy***	239	2.0%
9	Machine learning / algorithms***	229	1.9%
10	Language / speech / linguistics***	223	1.8%

Table B.8 South East

Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	2,342	16.7%
2	Climate / plants / crops / species	1,777	12.6%
3	Cells / cancer / proteins / infection	1,489	10.6%
4	Quantum / devices / reactions / batteries	1,329	9.5%
5	Galaxies / stars / universe / physics	814	5.8%
6	Brain / neurons / memory / neuroscience	552	3.9%
7	Equations / mathematics / geometry	509	3.6%
8	Machine learning / algorithms***	351	2.5%
9	Security / business / services / users***	279	2.0%
10	Music / games / sound***	272	1.9%

Table B.9 South West

Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	1,343	17.3%
2	Climate / plants / crops / species	1,188	15.3%
3	Quantum / devices / reactions / batteries	789	10.1%
4	Cells / cancer / proteins / infection	540	6.9%
5	Brain / neurons / memory / neuroscience	281	3.6%
6	Equations / mathematics / geometry	252	3.2%
7	Galaxies / stars / universe / physics	203	2.6%
8	Machine learning / algorithms***	199	2.6%
9	Manufacturing / steel / metals***	172	2.2%
10	Hydrogen / engines / fuels / vehicles***	170	2.2%

Table B.10 Wales

Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	761	21.3%
2	Climate / plants / crops / species	442	12.4%
3	Quantum / devices / reactions / batteries	383	10.7%
4	Cells / cancer / proteins / infection	212	5.9%
5	Galaxies / stars / universe / physics	159	4.5%
6	Brain / neurons / memory / neuroscience	153	4.3%
7	Manufacturing / steel / metals***	128	3.6%
8	Equations / mathematics / geometry	96	2.7%
9	Language / speech / linguistics***	88	2.5%
10	Plastics / waste / polymers / products***	78	2.2%

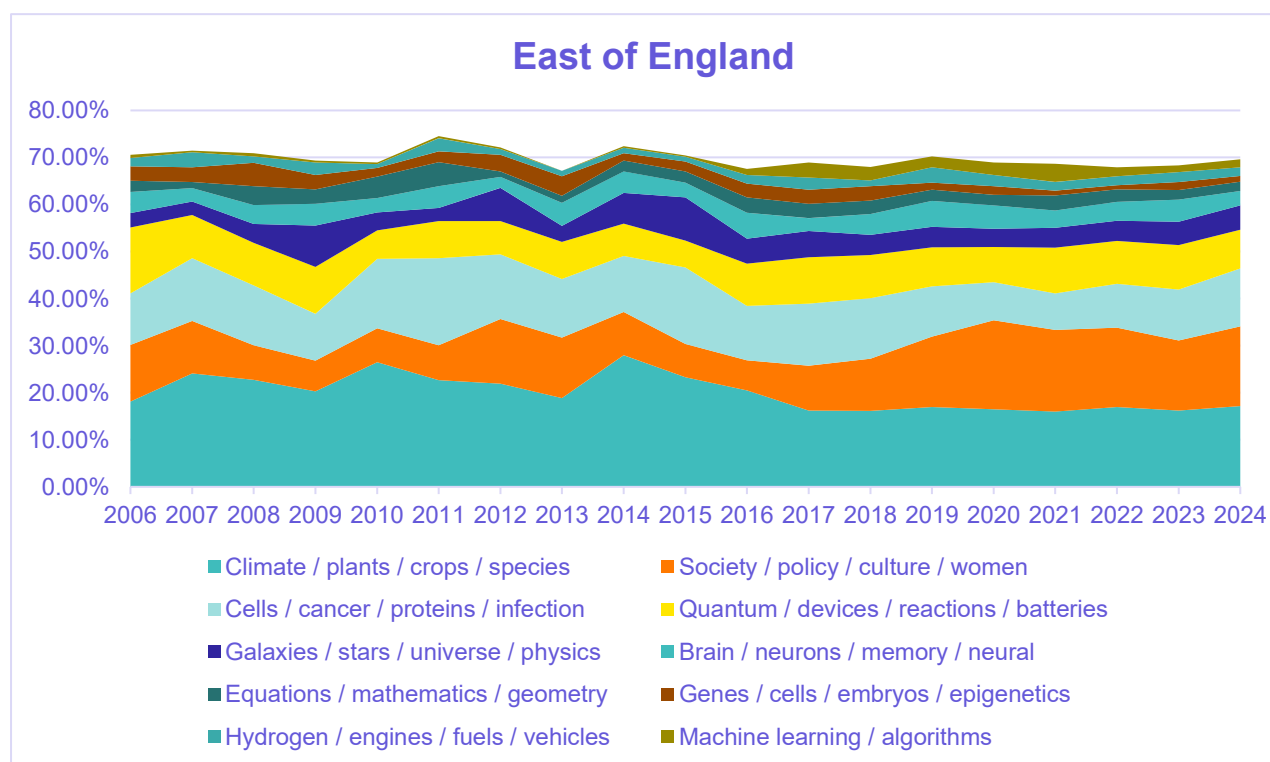
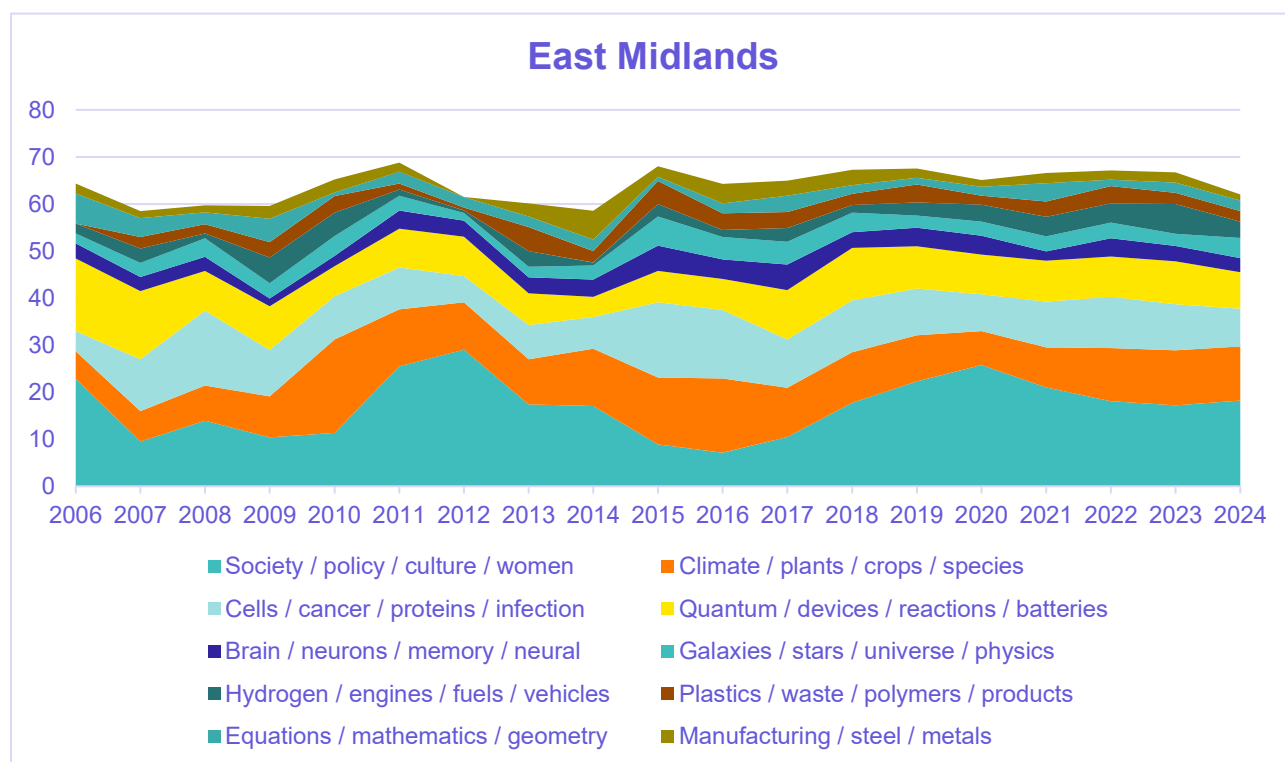
Table B.11 West Midlands

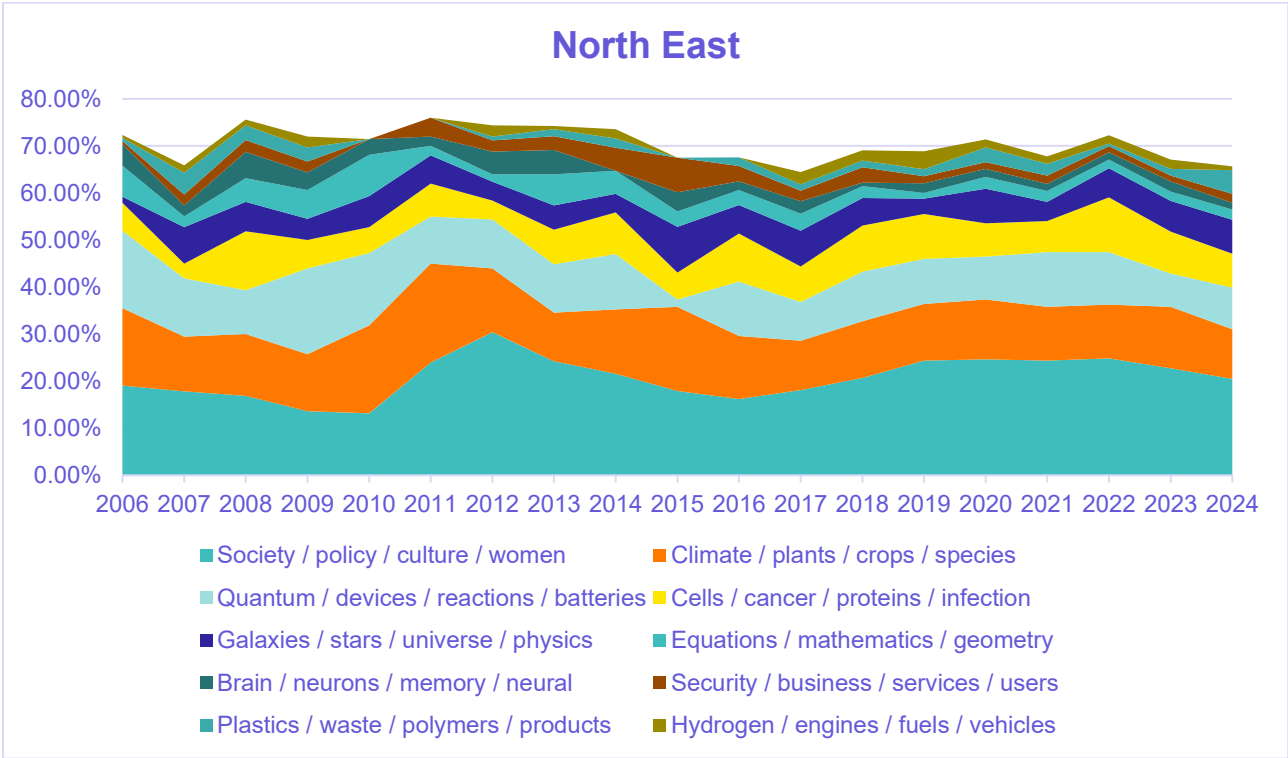
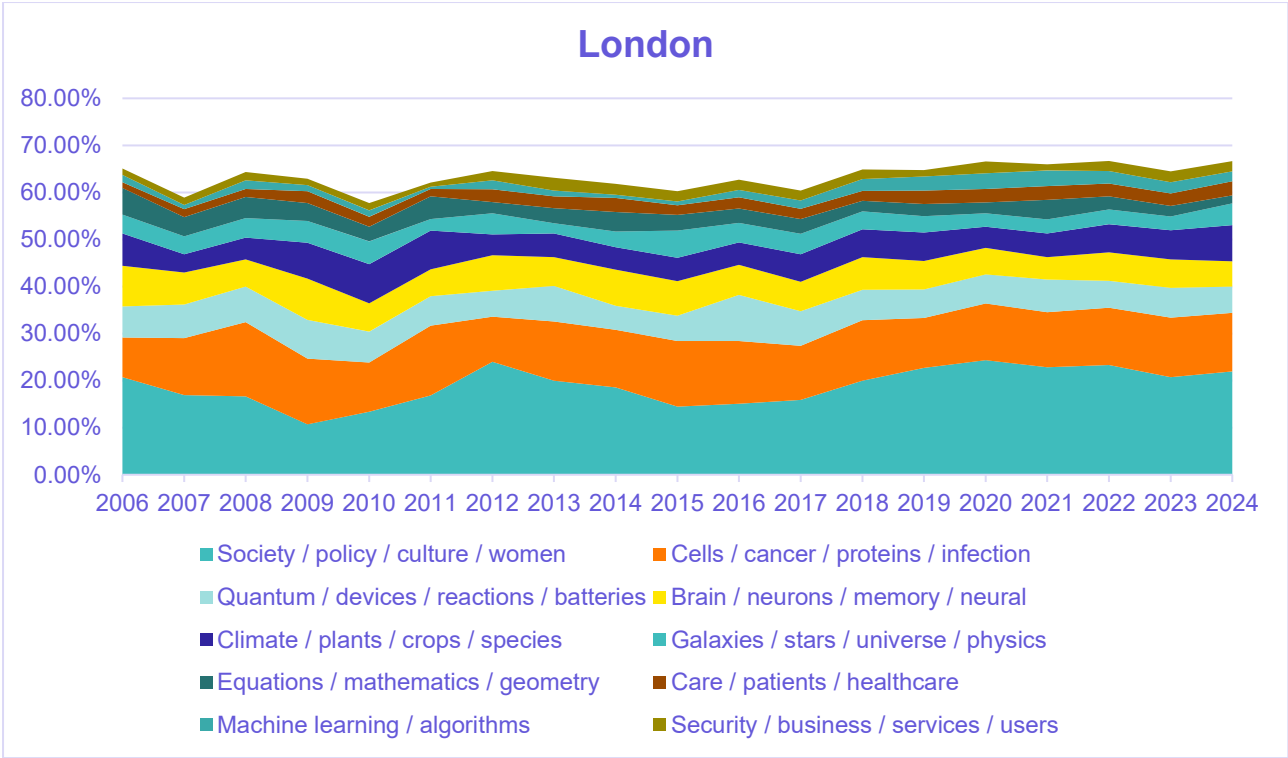
Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	1,160	18.7%
2	Cells / cancer / proteins / infection	753	12.1%
3	Quantum / devices / reactions / batteries	484	7.8%
4	Climate / plants / crops / species	478	7.7%
5	Equations / mathematics / geometry	300	4.8%
6	Galaxies / stars / universe / physics	292	4.7%
7	Brain / neurons / memory / neuroscience	172	2.8%
8	Manufacturing / steel / metals***	165	2.7%
9	Care / patients / healthcare***	136	2.2%
10	Machine learning / algorithms***	136	2.2%

Table B.11 Yorkshire and the Humber

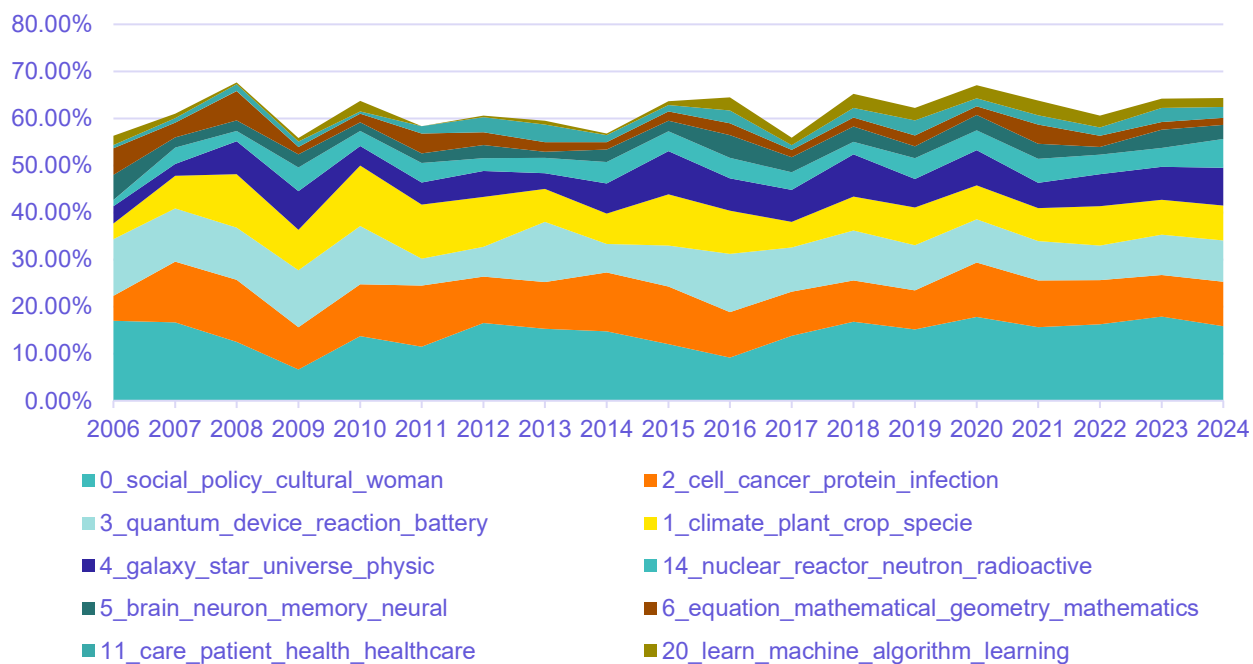
Rank	Topic	Topic records	Share within region
1	Society / policy / culture / women	1,419	16.3%
2	Climate / plants / crops / species	1,123	12.9%
3	Cells / cancer / proteins / infection	838	9.6%
4	Quantum / devices / reactions / batteries	803	9.2%
5	Equations / mathematics / geometry	243	2.8%
6	Galaxies / stars / universe / physics	237	2.7%
7	Manufacturing / steel / metals***	233	2.7%
8	Brain / neurons / memory / neuroscience	229	2.6%
9	Music / games / sound***	223	2.6%
10	Nuclear energy / reactors / neutrons / radioactivity***	209	2.4%

Appendix C. Thematic Evolution per Region

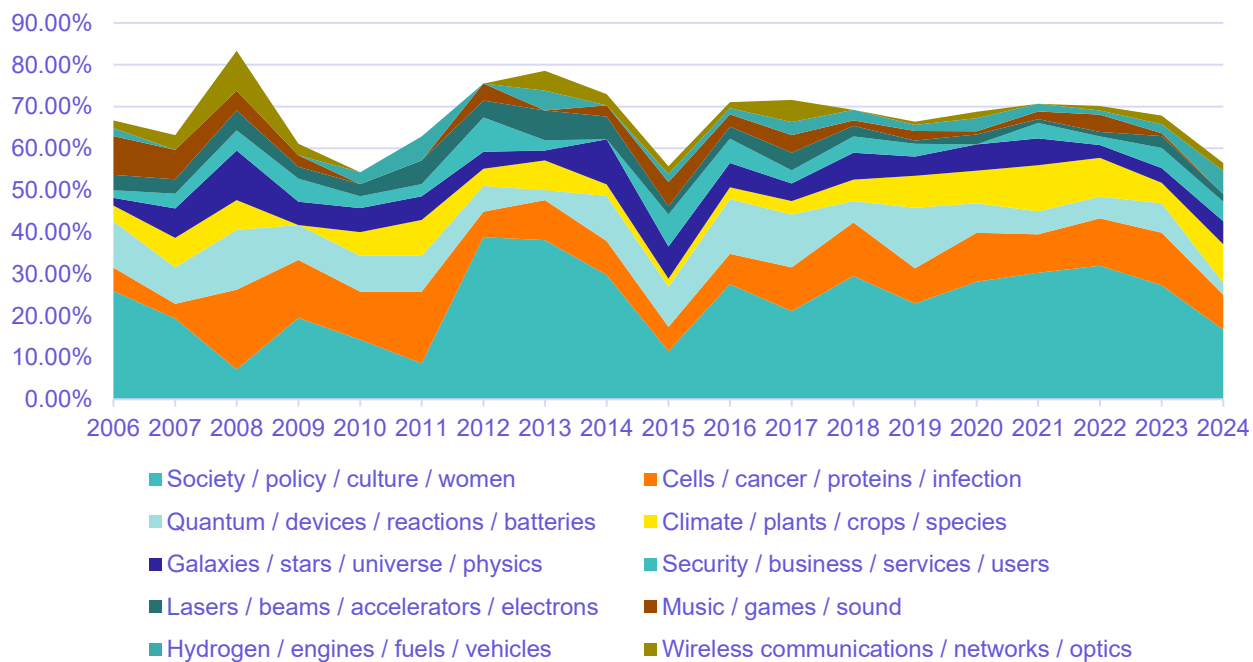


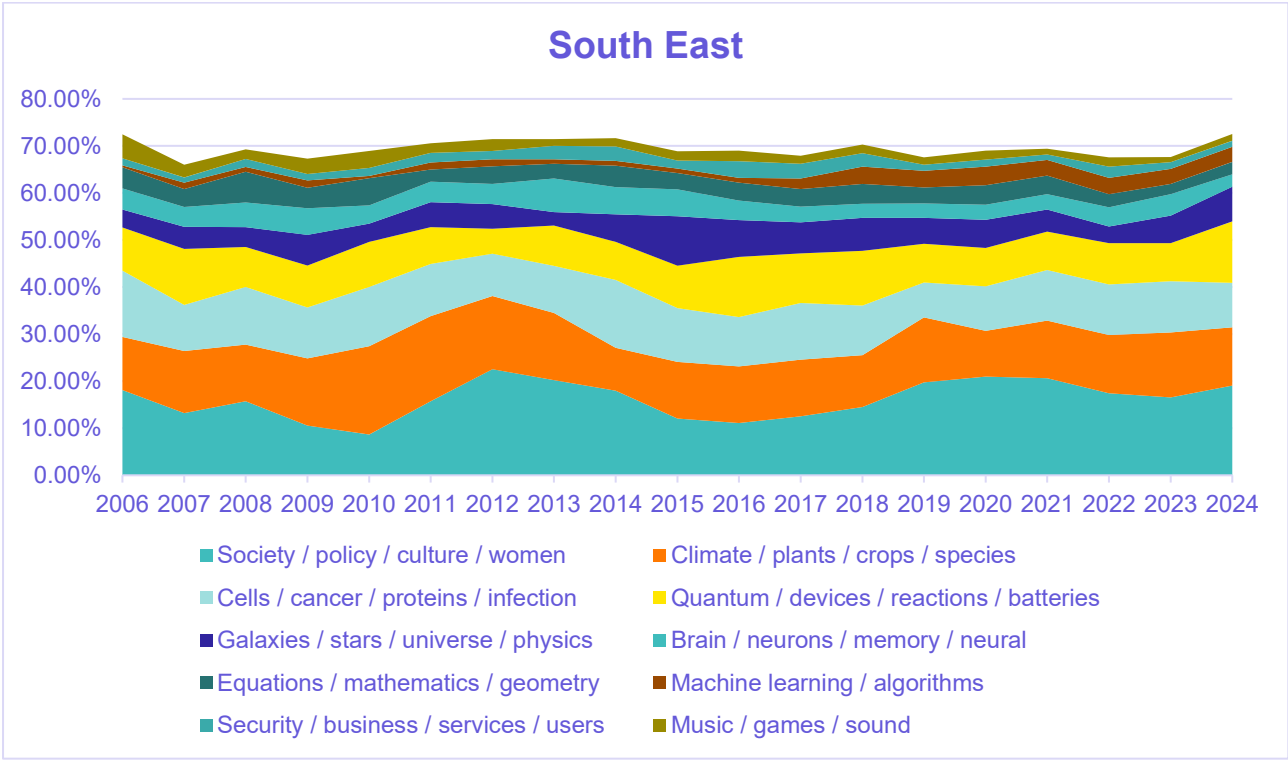
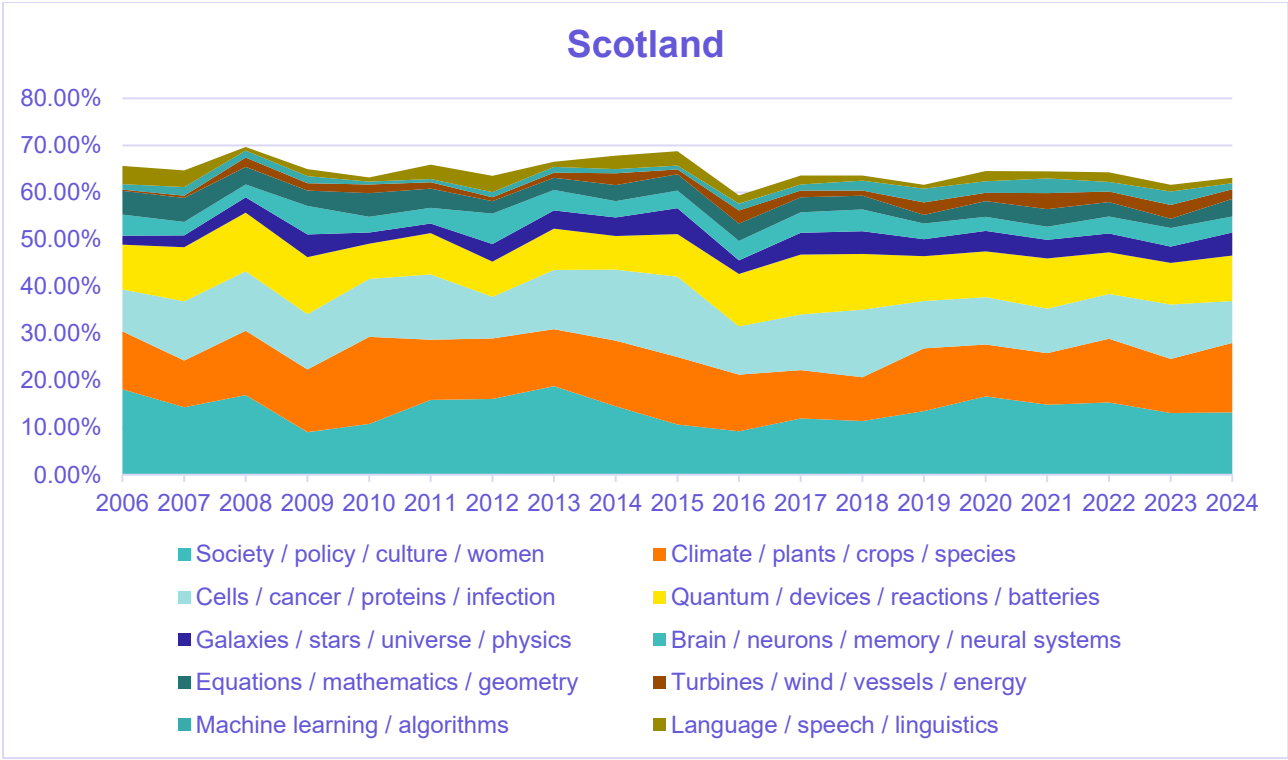


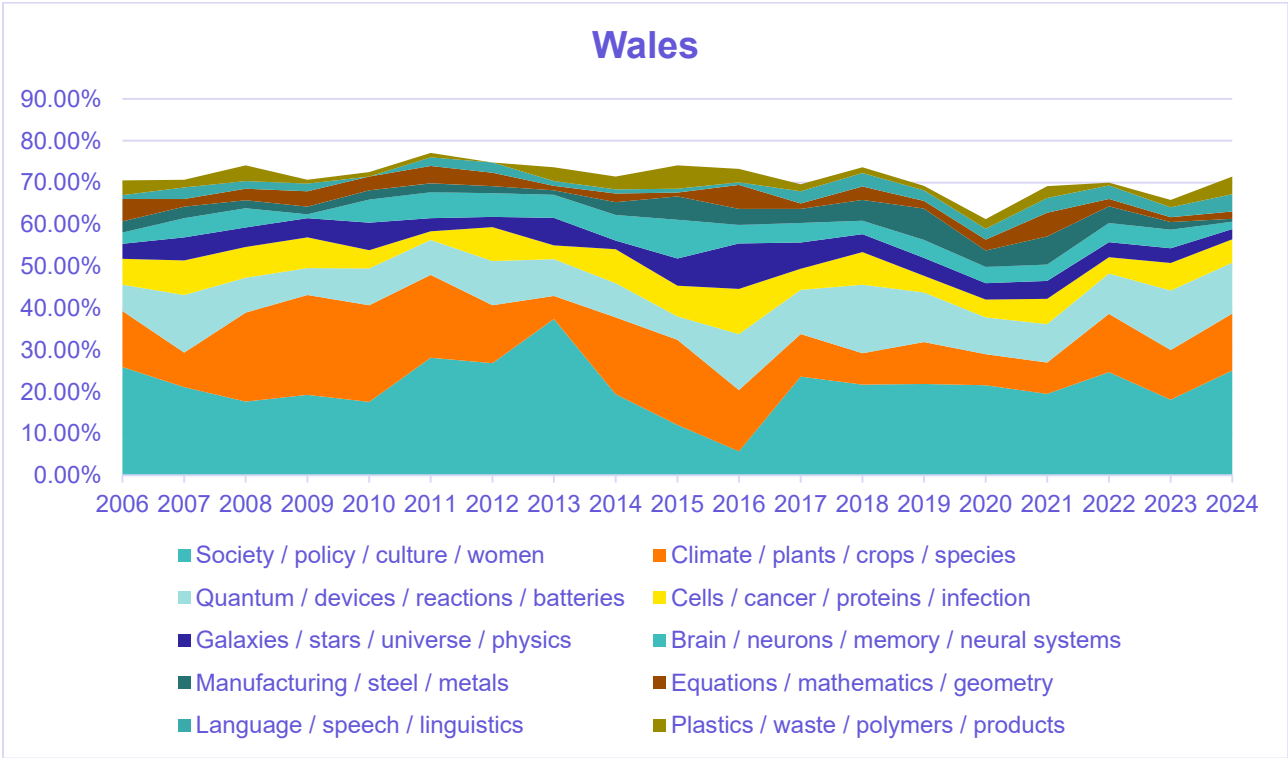
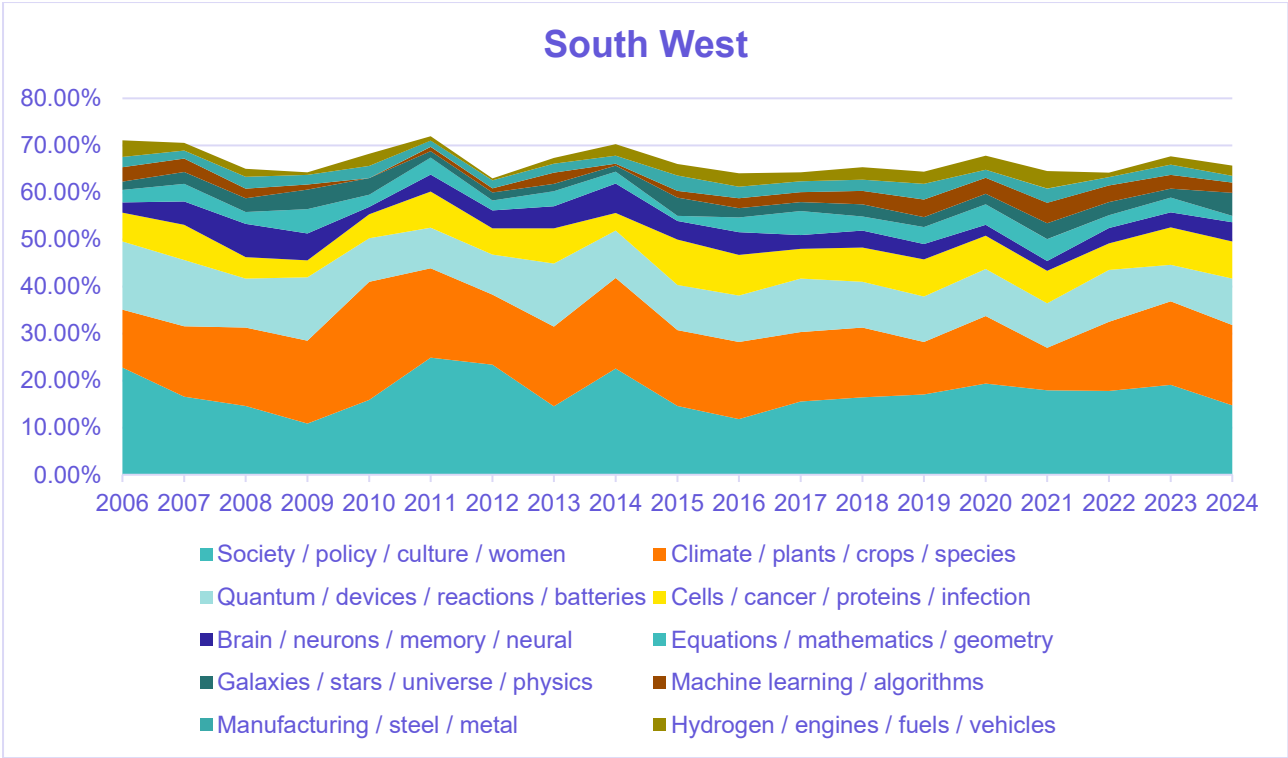
North West



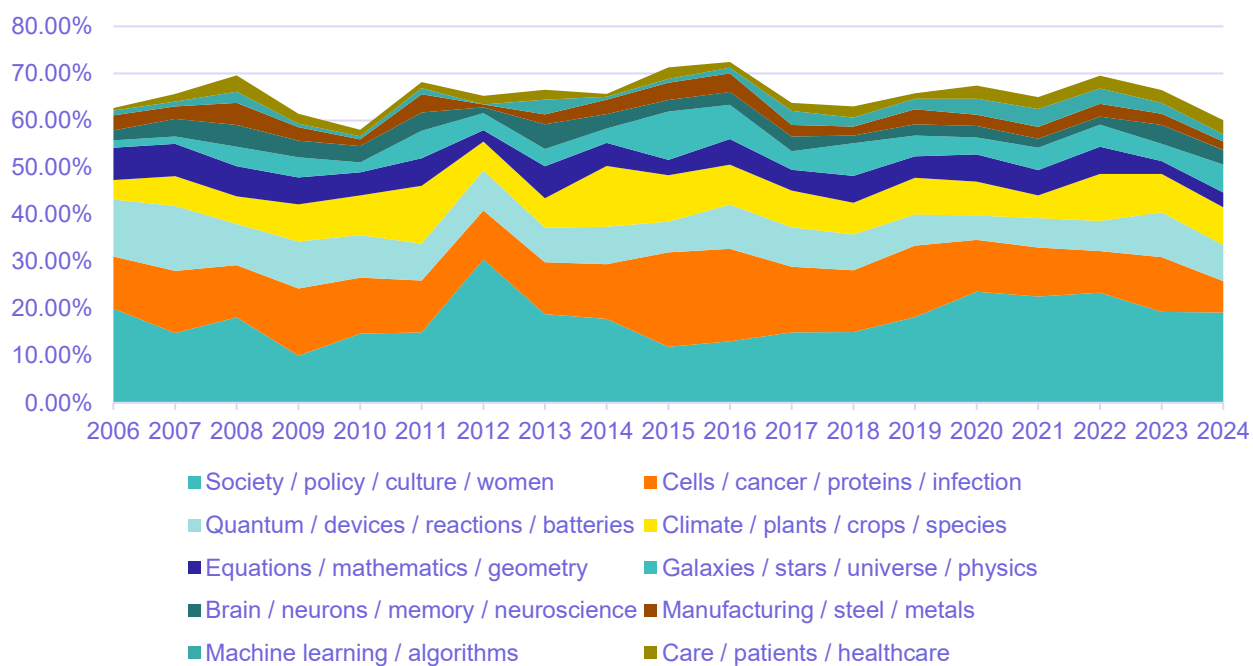
Northern Ireland



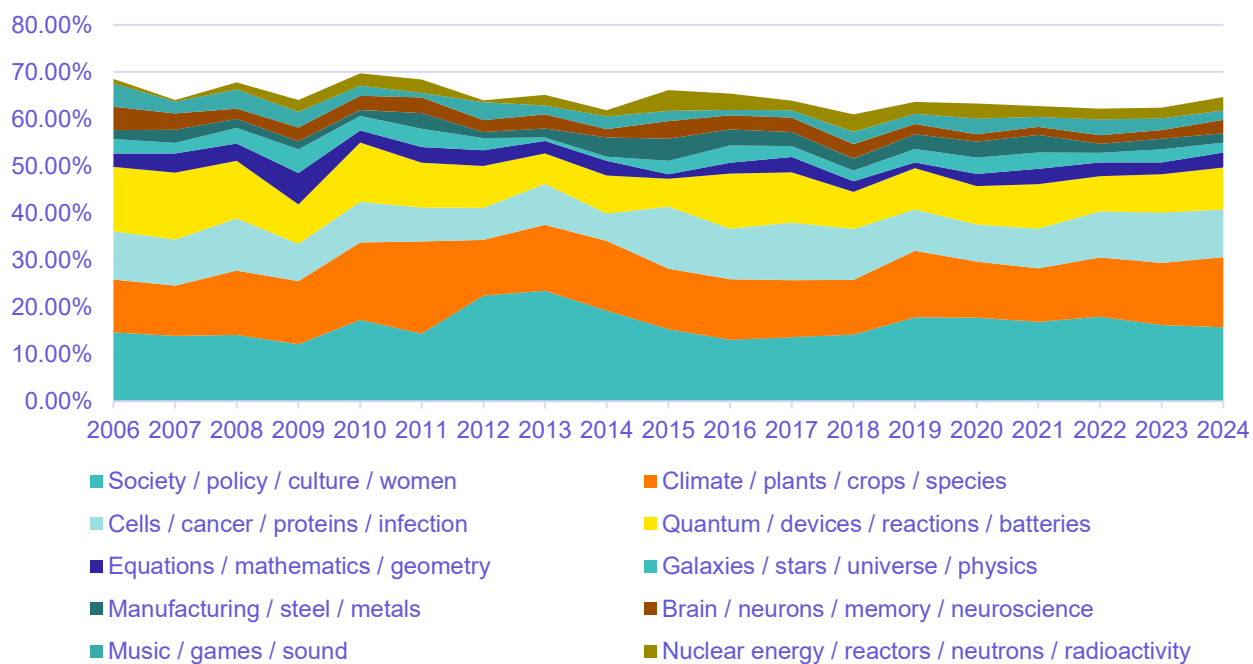




West Midlands



Yorkshire and the Humber





www.ircaucus.ac.uk

Email info@ircaucus.ac.uk Twitter [@IRCaucus](https://twitter.com/IRCaucus)

