The Use of Artificial Intelligence in Science, Technology, Engineering, and Medicine

A report prepared for The Royal Society

Glen Berman*, Jennifer Chubb**, and Kate Williams**

* Australian National University

** University of York

*** University of Melbourne

Submitted 9th May 2023

Contents

1 Executive summary

This report forms part of The Royal Society's *Disruptive Technology for Research* project which aims to understand the landscape of data-driven and AI based technologies across different fields of scientific research. In this report, we provide a **Taxonomy of Artificial Intelligence** related technologies for scientific research.

1.1 Summary of findings

Our analysis highlights the following features of the literature and publication analysis: There is substantial evidence that AI is used across the STEM areas. The use of AI related technologies in STEM is dominated by the physical sciences and medicine. Excluding computer science, the most prevalent use of AI technologies relates to engineering. The most prevalent use of AI in the health sciences relates to informatics and imaging. The most prevalent use of AI in the life sciences relates to neuroscience. Further, we find that the use of AI across STEM categories has increased over time, peaking in 2021.

Evidence of interdisciplinary work between the Computer Science fields and other fields is less than may have been anticipated. In particular there is a need for better cross-disciplinary communication and work. Inevitably, ethics and AI interpretability and transparency issues emerge as intertwined (though out of scope of this report). This poses an issue for training needs of professionals across all fields. A variety of use cases for top AI techniques and applications can be seen in impact case studies submitted to the UK REF[1] 2021 and are included in the appendix.

Finally, our taxonomy reveals that the AI technique most prominently used across all fields include Artificial Neural Networks (ANN) which is mentioned in almost all fields (73%), Deep Learning (63%), Internet of Things (63%), Machine Learning (63%) and Image recognition (52%). Under 50% of fields referenced Computer Vision, Convolutional Neural Networks (CNN), Robotics, and Big Data analytics.

1.2 Limitations and future research

The lack of a shared, concise definition of AI across the STEM fields makes identifying AI research challenging. In addition, classification systems of scientific fields and subfields are imperfect attempts to

[1] The REF is the UK's system for assessing the quality of research in UK higher education institutions. https://www.ref.ac.uk/

capture the boundaries around different approaches and activities. As such, in this report we have used a mix of top-down and bottom-up methods to surface trends in the adoption of AI techniques in the different fields, and use different sources of data across the report.

Further research would rigorously examine these data in more detail. Moreover, an empirical exploration of researchers' attitudes and experiences of using these techniques within AI research, delving into the interdisciplinary aspects focusing on the interaction between and across the social sciences, arts and humanities would be helpful. As noted in the summary, ethics emerge as intertwined with and inseparable from the adoption of AI in STEM particularly in areas relating to health and medicine. Further research which seeks to address these core aspects would enable a deeper understanding of the issues of responsible AI innovation.

This short term project provides only a snapshot of the current state of the art in this area. There are no comparators which can be made at this stage in terms of the use of AI in research given the project scope (for instance; with respect to interdisciplinarity, conclusions cannot be made vis a vis AI research and non AI research practices). Further research can address this.

2 Introduction

2.1 Aims of this report

The Royal Society's *Disruptive Technology for Research* project aims to understand the landscape of data-driven and artificial intelligence-based technologies (AI) across different fields of scientific research. This document first provides a literature review of AI use in Science, Technology, Engineering and Medicine (STEM). Second, using data from the Scopus database, it addresses the following questions:

1. Which STEM fields are using AI technologies?
2. What proportion of fields are using AI?
3. What are the main subfields using AI?
4. How have the fields using AI technologies changed over time?
5. How are different fields correlated with each other?
6. What technologies are prominent within fields?
7. What are the characteristics of the AI research area?

Third, it provides a taxonomy that sets out the top fields and subfields in AI and the main AI techniques they utilise. We do this drawing on themes from the literature review and example use cases from the publicly available UK Research Excellence Framework (REF) Impact Case studies (ICS). This research aims to enable end users to have an overview of the current state of commercialisation and application for AI related inventions in scientific research - adding context to the challenges, opportunities and Royal Society's recommendations.

2.2 Report structure

The report is structured as follows:

- **Section 3**: A literature review of AI use in the STEM fields, reviewing recent evidence in the published literature from the period 2018 to 2023;
- **Section 4**: A summary of trends in AI publications across the STEM fields, focused on visualising the extent to which different fields are enrolling AI techniques in their research; and
- **Section 5**: A high-level taxonomy that outlines applications of AI techniques across research in the STEM fields.

Alongside the above sections, we provide a methodological note on our analysis in **Section 2.3**.

2.3 Methodological note

In this report we use 'AI' as a broad term, covering all efforts aiming to replicate and extend human capabilities for intelligence and reasoning in machines. Since the founding of the AI field at the 1956 Dartmouth Summer Research Project on AI, many different techniques have been invented and studied in pursuit of this goal. Many of these techniques have developed into their own subfields within Computer Science, such as expert systems and symbolic reasoning. Machine Learning (ML) is one such technique, as is the dominant ML paradigm, Deep Learning (DL).

Since the rise of Big Data and the advent of cost-effective parallel computing, ML and DL techniques have achieved remarkable successes in a wide variety of research and industry domains—so much so, that in modern parlance 'AI' is often treated as synonymous with ML. Reflecting this, this report—and the Disruptive Technology for Research project more generally—focuses on the use of ML and DL techniques in the sciences.

The lack of a shared, concise definition of AI across the STEM fields makes identifying AI research challenging. As such, in this report we use a mix of top-down and bottom-up methods to surface trends in the adoption of AI techniques in the different fields, and utilise different sources of data across the report.

First, in the Literature Review (**Section 3**), we primarily rely on a bottom-up approach: we search through the academic literature on various STEM fields for reviews of the use of AI technologies and synthesise trends found in these reviews. The literature review thus presents a perspective of AI use from within each field, allowing for divergence between fields in terms of what constitutes 'AI', and the language used to describe specific AI techniques.

Second, in the analysis of trends in AI publications (**Section 4**), we primarily rely on a top-down approach: we use Elsevier's dataset of AI publications, a large dataset of AI publications across all fields, which Elsevier has prepared through expert analysis of keywords associated with AI research (there are 800+ keywords used to curate the dataset). This dataset enables us to identify a very wide range of AI publications, and extract bibliometric data associated with them. As the focus of this report is the current use of AI techniques in STEM research, we extract and analyse all AI publications for the years 2018 - 2022.

Third, we supplement the data described above with insights from Impact Case Studies (ICS) prepared in the UK for the Research Excellence Framework (REF). This is the means by which research is assessed in UK research institutions and funds are allocated. Among other purposes, REF states it seeks to provide benchmarking and accountability around impact beyond academia. ICS outline the underpinning research and provide a narrative about the resulting impact in society in terms of reach and significance. The ICS therefore provide a snapshot of where AI is used in research and the resulting applications and public reach. Using selected keywords from the Elsevier AI dataset we identified 202 Impact Case Studies that detail the use of AI technologies. Then we filtered to exclude those outside of the remit (e.g. social sciences). This revealed a total of 130 ICS referencing AI. Through desk based research, we thematically analysed reference to AI in Science where the underpinning research utilised

The analysis itself took place in three stages. The first stage was the comprehensive literature review of academic work on AI technologies and research practice. The second was the analysis of publication data from Elsevier's Scopus to support a field-level overview of the use of AI in scientific research. The third stage was analysis of how AI technology features in research practice, dissemination and impact by using the UK REF ICS database. Together, the literature review, publication data and impact case studies were used to provide field-level insights into the use of AI technologies in research.

2.3.1 Additional notes on the literature review

To structure the literature review, the subject categorisations from Elsevier's All Science Journal Classification Codes (ASJC)[2] were used, as shown in **Table 1.** Given the focus of this review on STEM subjects, the subject area of 'Social Sciences and Humanities' was excluded.

Literature searches were conducted by inputting AI related keywords (e.g. artificial intelligence, machine learning, deep learning) in combination with the subject area classifications listed below. The search concentrated on the most recent articles (i.e. limited to 2018-2023) that provided systematic or comprehensive reviews on the field, in order to gain oversight into the key trends within each field.

---

[2] A list of subject areas can be found at https://service.elsevier.com/app/answers/detail/a_id/12007/supporthub/scopus/, while details of 'Subject Area Classifications' can be found at https://service.elsevier.com/app/answers/detail/a_id/15181/supporthub/scopus/

Table 1 Subject categorisation using All Science Journal Classification Codes.

Subject Area	Subject Area Classifications
Physical Sciences	Chemical Engineering Chemistry Computer Science Earth and Planetary Sciences Energy Engineering Environmental Science Material Science Mathematics Physics and Astronomy Multidisciplinary
Health Sciences	Medicine Nursing Veterinary Dentistry Health Professions Multidisciplinary
Life Sciences	Agricultural and Biological Sciences Biochemistry, Genetics and Molecular Biology Immunology and Microbiology Neuroscience Pharmacology, Toxicology and Pharmaceutics Multidisciplinary

2.3.2 Trends in AI publications across the STEM fields

To analyse trends in AI publications across the STEM fields we undertook bibliometric analysis of all AI publications in the Elsevier AI dataset between 2018 and 2022. Our primary focus here was on comparing the volume and content of AI research across different fields. To do so we began by disaggregating AI publications into their STEM fields, through use of the ASJC codes. Each publication is assigned by Elsevier multiple ASJC codes, reflecting the thematic focus of the journal in which the publication is contained. ASJC codes are hierarchical. A snapshot of the hierarchy is shown in **Table 2**, which reports two of the fields, subfields, and ASJC codes that fall within the 'physical sciences' category. As can be seen, ASJC codes correspond to subfields.

By default, when Elsevier calculates category, field, or subfield level data on publications, every publication is counted once for each of its ASJC codes. As such, a publication with five ASJC codes will be counted in a breakdown of the AI dataset by 'field' five times, which can lead to substantial distortions in the analysis. To avoid this, we fractionalise all publications according to each of their ASJC codes. If a publication has five ASJC codes, then it is counted as 0.2 of a publication for each code. This ensures that our analysis of the distribution of publications across categories, fields, and subfields, does not inflate the total number of publications. Additionally, because a publication may be associated with multiple ASJC codes from different fields or categories, our fractionalisation approach ensures that interdisciplinarity is somewhat accounted for in our analysis: a publication that is associated with four ASJC codes from 'Earth and Planetary Sciences' subfields and one ASJC code from a 'Physics and Astronomy' subfield will be counted as 0.8 of an 'Earth and Planetary Sciences' publication and 0.2 of a 'Physics and Astronomy' publication. Visualisations 1, 3, and 4 in **Section 4** are produced through analysis of this fractionalised data.

Table 2 Snapshot of ASJC code hierarchy

Category	**Field**	**Subfield**	**ASJC code**
Physical sciences	Earth and Planetary Sciences	Atmospheric Science	1902
Physical sciences	Earth and Planetary Sciences	Computers in Earth Sciences	1903
Physical sciences	Earth and Planetary Sciences	Earth and Planetary Sciences (miscellaneous)	1901
Physical sciences	Earth and Planetary Sciences	Earth-Surface Processes	1904
Physical sciences	Earth and Planetary Sciences	Economic Geology	1905
Physical sciences	Earth and Planetary Sciences	General Earth and Planetary Sciences	1900
Physical sciences	Earth and Planetary Sciences	Geochemistry and Petrology	1906
Physical sciences	Earth and Planetary Sciences	Geology	1907
Physical sciences	Earth and Planetary Sciences	Geophysics	1908
Physical sciences	Earth and Planetary Sciences	Geotechnical Engineering and Engineering Geology	1909
Physical sciences	Earth and Planetary Sciences	Oceanography	1910
Physical sciences	Earth and Planetary Sciences	Palaeontology	1911
Physical sciences	Earth and Planetary Sciences	Space and Planetary Science	1912
Physical sciences	Earth and Planetary Sciences	Stratigraphy	1913
Physical sciences	Physics and Astronomy	Acoustics and Ultrasonics	3102
Physical sciences	Physics and Astronomy	Astronomy and Astrophysics	3103
Physical sciences	Physics and Astronomy	Atomic and Molecular Physics, and Optics	3107
Physical sciences	Physics and Astronomy	Condensed Matter Physics	3104
Physical sciences	Physics and Astronomy	General Physics and Astronomy	3100
Physical sciences	Physics and Astronomy	Instrumentation	3105
Physical sciences	Physics and Astronomy	Nuclear and High Energy Physics	3106
Physical sciences	Physics and Astronomy	Physics and Astronomy (miscellaneous)	3101
Physical sciences	Physics and Astronomy	Radiation	3108
Physical sciences	Physics and Astronomy	Statistical and Nonlinear Physics	3109
Physical sciences	Physics and Astronomy	Surfaces and Interfaces	3110

A limitation of our reliance on Elsevier's AI publications dataset is that this data is insufficient to undertake comparative analysis of AI publications versus non-AI publications. Given the breadth of this report, focusing on AI publications in all STEM fields, such a comparative analysis would effectively require a dataset of *all* publications in all STEM fields – which, while theoretically possible to access, was well beyond the scope of this project. As such, to provide some intuition as to what proportion of the STEM fields are making use of AI techniques in their research (visualisation 2 in **Section 4**), we utilise Elsevier's data on all publications in all STEM fields. This data is unfractionalised, but nonetheless provides some indication of the extent to which different STEM fields are making use of AI fields.

To assess how different STEM fields collaborate when making use of AI technologies (visualisation 5 in **Section 4**), we again made use of ASJC codes. Here, once publications were fractionalised, we considered correlations between the co-occurrence of fields in the AI dataset. In other words, we considered the likelihood, in a given publication, of two ASJC field codes co-occurring. To do so we calculated the Pearson correlation coefficient across the fields, which we report in a correlation matrix. We note that this analysis cannot shed light on whether collaboration between STEM fields when making use of AI technologies differs from collaboration between STEM fields generally (as doing so would require a baseline comparator which was outside the scope of this project to develop).

Finally, we attempt to provide some insights as to the underlying AI technologies that are being relied on in different STEM fields. Unfortunately, in this instance our analysis is subject to the limitations of bibliometric data. Despite substantial experimentation with ASJC codes, as well as other classification schema provided by Elsevier, we found that no schema offered insights at the relevant level of detail to be of use to this report. For example, Elsevier's Topic Clusters and Topic name schema do both have categories called 'Deep learning', however these categories are used so broadly that they provide little useful insight as to what specific deep learning techniques are in use across the STEM fields. The fundamental challenge here is that AI techniques are developing and changing at a far faster rate than bibliometric classification schema. In future work we may be able to provide some insights into these questions through topic modelling of publication abstracts, which would enable us to generate our own classification schema, however this was outside the scope of this project (and, of course, topic modelling introduces its own limitations). To address this limitation, we used a bottom-up method of identifying relevant AI techniques, as shown below.

2.3.3 Taxonomy

For the taxonomy, we produced an ordered list of fields as a percentage of the three overarching categories (Health, Life, Physical Sciences), and then identified the top three subfields per field. Next, to identify AI techniques for each field, we used a bottom up, grounded approach and conducted a coding exercise from which to generate techniques largely inductively. From this, we state the main topics (techniques) described in literature review and the publicly available UK REF ICS at field level. We developed categories and codes from the data, and analytical memos were used between coding and writing. Pre-existing conceptualisations of AI techniques were known to the researchers, including the most prevalent AI techniques such as Machine Learning, Natural Language Processing and Artificial Neural Networks, clearly shown either in the title of the literature or ICS classifications. As such, the approach can be said to involve a level of inductive coding. Coders summarised case study content pertaining to each code, for example by listing examples of AI techniques. The most prevalent were included in the simple taxonomy.

2.3.4. Limitations across this report

Across this report, our focus has been on assessing the breadth of AI techniques used in the STEM fields, rather than the depth. As such, our methods necessarily focus on high-level analysis of trends across fields, rather than on detailed analysis of individual publications, or fields. Additionally, we make no attempt to distinguish influential AI research from non-influential research. Given the volume of AI research produced, there may be significant divergences in trends (e.g. in collaboration, in topic focus) expressed at the level of all STEM fields, compared to trends expressed within publications in the STEM fields that are highly cited, or from highly esteemed institutions. Our research is unable to address this.

Further, due to the significant time and resource constraints informing this report, we have made use of pre-existing datasets and topic schema (specifically, Elsevier's AI dataset and the ASJC schema). While use of these sources and schema is widespread in bibliometric analyses, they are nonetheless not without their flaws. Most relevant to this research project, evaluating the accuracy or correspondence between allocation of ASJC codes and the substantive content of research publications was outside of scope. As such, while ASJC codes are generally considered reliable, we are unable to validate whether the trends we have identified (particularly in **Section 4**) are manifestations of substantive trends in research, or merely manifestations of superficial quirks in the classification schema.

3 Literature review

3.1 Introduction to the literature review

AI is ubiquitous in society - it is part of our everyday lives. Crucially, AI, in particular ML, is being used across all fields and sectors from the creative arts through to cybersecurity and health diagnostics. In recent years, there has been a growing concern over the ethical implications of ML. This context is particularly important when exploring the use of AI in STEM where much of the literature attempts to tackle both of these aspects though outside of the scope of this review.

Applications of AI can be found across all STEM fields, with concentration in digital health, medicine and dentistry, environment, materials science, engineering, computational and mathematical sciences, robotics, quantum information technologies, health informatics, computer aided design, nurse and health care professional education and training, sustainability, dynamical and control systems engineering, safety critical systems, genetics, agriculture, energy, conservation and efficiency, data analytics, mechanical engineering, smart innovation, physics and astronomy, and computer and information science. Crucially, the application of AI in STEM looks to consider its application from a range of perspectives from policy to practice.

Within STEM, AI is widely adopted in fields that have (1) substantial data infrastructure and (2) a history of using computer models. For adoption of AI techniques these two requirements are mutually reinforcing: data infrastructure means there is surfeit of data for training models, and a history of using modelling means there are scientists who feel comfortable translating and adopting AI techniques to their field.

3.1.1 Approach to the literature

To produce this summary of the key themes identified in the literature, the following subject categorisations were used, as given in Table 1 above.:

- **Physical sciences:** Chemical Engineering, Chemistry, Computer Science, Earth and Planetary Sciences, Energy, Engineering, Environmental Science, Material Science, Mathematics, Physics and Astronomy.
- **Health sciences:** Medicine, Nursing, Veterinary, Dentistry, Health Professions.

- **Life sciences:** Agricultural and Biological Sciences, Biochemistry, Genetics and Molecular Biology, Immunology and Microbiology, Neuroscience, Pharmacology, Toxicology and Pharmaceutics.

Literature searches were conducted by inputting AI related keywords (e.g. artificial intelligence, machine learning, deep learning) in combination with subject areas listed above. The focus was on recent systematic or comprehensive reviews in order to gain oversight into the key trends within each field.

3.2 Summary of key themes identified in the literature

This section outlines the domains of use of AI across the subfields of STEM. Each area describes the ways in which AI is deployed, the benefits and key challenges highlighted in the literature.

3.2.1 Physical sciences

AI and the physical sciences are deeply entwined. Indeed, the physical sciences have a long history of deploying AI techniques to advance their research objectives. For example, expert systems, an AI technique which was actively researched throughout the 1970s and 1980s and aimed to replicate the decision-making processes of a domain expert, were adopted in chemical engineering and chemistry. Today, many subdomains in the physical sciences make substantial use of AI techniques to model natural and anthropogenic phenomena—from natural disasters to synthetic materials—and to help extract information from rapidly accumulating data streams, such as the data generated by experiments at the Large Hadron Collider. Meanwhile, subdomains in the physical sciences are also helping advance AI techniques, both by creating new challenges for AI techniques to address and by providing domain expertise to enable development of new classes of ML models. Of particular relevance to both the AI field and the physical sciences is the development of hybrid physics-based ML models—models that learn both from data and from domain-specific principles—as these models have demonstrated significant potential in addressing core AI challenges of model interpretability, robustness, and reliability.

Within the physical sciences, the fields of computer science and mathematics are particularly closely associated with the AI field. ML, in particular, is often considered a subfield of computer science. Current research focuses across computer science include the development of new ML modelling approaches, including advanced DL and generative ML, and the application of these approaches to highly generalisable tasks, such as natural language processing, image processing and computer vision, game playing, and model explanation and interpretation. Additionally, within applied subfields of computer science, particularly cybersecurity, video game development, and software engineering, a major focus has been the integration of ML techniques into workflows and data processing. Current research focuses across mathematics, meanwhile, include addressing some of the fundamental research challenges in AI: identifying optimal model training processes, model interpretability and robustness, and quantifying the extent of model generalisability. These challenges reflect, in particular, the ongoing search for a mathematical foundation for DL, to complement the growing empirical evidence of DL's effectiveness.

The physical sciences have been particularly receptive to the current wave of ML and DL research. In part, this is because the development of ML models to ingest and identify patterns in very large datasets, and the maturation of open-source software to build ML models, has coincided with explosive growth in data generated by various physical science fields. In high energy particle physics, for example, the Large Hadron Collider is producing a vast volume of data, and the search for new particles requires the ability to detect minute signals within this data. Similarly, in astronomy data generated by the Event Horizon Telescope is of such a scale that information extraction is only possible with the support of custom-built ML models. Comparable advances in data generation can be found in the fields of environmental science, energy, earth and planetary sciences, chemistry, and chemical engineering. In all these fields significant growth in the array of sensors used to monitor and observe physical phenomena has resulted in comparably significant growth in the volume of data generated, creating clear use cases for the deployment of ML techniques.

An additional factor in the enthusiastic adoption of ML and DL research in the physical sciences is the extant centrality of computer-based modelling to many fields. Pre-dating the current wave of ML, modelling was already used in many physical sciences to complement and extend physical experiments, and to manage industrial processes. In chemical engineering, models are critical to understanding chemical reactions, and are used to monitor chemical manufacturing processes. Similarly, a core focus of research in the Earth and planetary sciences has been the development of models of the climate, water systems, and mineral deposits. ML and DL techniques are extending the capacity of the physical sciences to model complex phenomena, particularly in instances where many variables interact. In chemical engineering and material science, ML and DL techniques have thus been applied to the problems of forward and inverse modelling of molecular structures, enabling more efficient and accurate prediction of the properties of a given molecular structure and the prediction of new molecular structures that are likely to produce specific desirable properties. In engineering, energy, chemistry, and environmental sciences, ML and DL techniques have been applied to problems of fault detection in industrial processing (e.g. chemical synthesising) and health monitoring of industrial and natural systems (e.g. structural health of infrastructure, health of waterways).

Three challenges to further use of AI techniques in the physical sciences stand out. First, while many physical sciences have strong data infrastructures in place, there are data gaps relating to publicly accessible datasets for model training, benchmark datasets for model evaluation, and benchmarks for the collection and curation of new datasets. Second, although researchers have shown interest in adopting ML and DL techniques, field-specific training in the appropriate use of these techniques is still under development. Finally, as ML and DL techniques are deployed in the physical sciences, limitations inherent in these techniques are also surfaced. Limitations of particular concern to the physical sciences are: scientific interpretation of model outputs; determination of model generalisability; and, most importantly, the potential for ML models to produce outputs that do not accord with known principles of natural phenomena. This final limitation has spurred the development of hybrid physics-informed ML models, which attempt to hard code into the neural network relevant laws of physics, to ensure model compliance. In the Earth and planetary sciences, energy, engineering, and physics and astronomy, physics-informed models are a significant area of ongoing research and collaboration with computer scientists.

3.2.2 Health sciences

AI is rapidly spreading across healthcare. The literature suggests that healthcare professionals prefer to use AI as a tool which must be used to complement human judgement and exposes concerns about AI explainability and transparency with frequent references to AI as a 'black box' technology. This is reinforced by frequent references to the implications of algorithmic historical bias and discriminatory data. Indeed, much of the literature on healthcare and AI use reveals ethics and AI to be deeply intertwined requiring patient and public involvement in the research process. Further, the literature reveals concerns over a lack of communication and understanding between clinicians and health professionals, and AI scientists.

Across the field of health sciences, the main subdomains where AI is used includes: mental health and psychiatry, pharmaceutical industries, physiology, neuroscience and neurology, infectious diseases, biotechnology, genetics, oncology, stem cell research, radiography, perinatal and gynaecology. To do this, fields rely on big data, machine learning, deep learning, and natural language processing as examples of AI-based technologies.

As a result of the COVID-19 pandemic, perhaps unsurprisingly, there has been an explosion of research about AI and digital health in the area of infectious disease and clinical management. Beyond this, a high level analysis of the ways in which AI is used in health shows that most commonly the areas of application are explored through the lens of disease detection, disease prediction, modelling, diagnosis, treatment, discovery, repair, data visualisation, surgery, training, robotics, precision, treatment, risk stratification, clinical management and decision making and health care ‘futures’ such as considerations of major grand challenges including risk prevention of infectious disease, ageing population and obesity.

AI offers huge potential for improving healthcare in supporting clinical decision making and new techniques for diagnostics and detection as well as clinical evaluation. However, its deployment is often opaque, and adoption can be ‘inhibited by the use of ‘black box’ AI systems where it is not understood why AI is an effective technique or how far it protects the rights of patients to confidentiality, consent and autonomy. Indeed, the rapid integration of AI in health has occurred with little communication between those developing it, computer scientists and doctors where, for instance, historical and algorithmic bias may widen healthcare inequalities.

3.2.3 Life sciences

AI and the life sciences have benefited from strongly intertwined progress. The literature shows two key areas. First, in the life sciences around agriculture and the food industry, advancements in the Internet of Things (IoT) have provided a range of sophisticated sensors and physical devices that in turn have provided a wealth of data that can be analysed using supervised and unsupervised AI models. The insights derived from these models are used in creating automated ‘smart machines’ or in informing human interventions. Second, in the life sciences oriented towards improving health outcomes and knowledge on biological and molecular systems, advancements in computing capability and algorithms have allowed the integration of big data, AI, and multi-scale modelling techniques to improve all stages of the process from research to clinical practice.

Across the field of the life sciences, the main subdomains include a focus on; food security, climate change, agriculture, farming, e-waste, water, energy, oceans, fish farming, analysis of power usage, soil management, water management, biochemistry, molecular biology including genetics, genomics, drug interactions and discoveries, immunology and neuroscience as well as modelling in pharmacology and toxicology. The sector particularly relies on the IoT, big data, neural networks, modelling and image recognition.

A high-level analysis of the ways in which AI is used in the life sciences shows the areas of application relate to autonomous farming, crop management, water management, soil management, livestock management, laboratory processes and care pathways, imaging and laboratory services, diagnosis and disease monitoring, patient eligibility, genomic analysis, diagnosis, prediction and treatment for psychiatric, neural and developmental disorders as well as drug development.

Overall, the life sciences are faced with two key challenges, which differ depending on the goals and processes of the subfield. Within the life sciences focused on agriculture and food, data is plentiful thanks to the relatively low cost of sensors and devices, however the use of advanced technologies are localised to highly developed countries, particularly large factory farms in North America and Europe, while making a livelihood remains extremely difficult for most farmers around the world. On the other hand, a challenge for the life sciences oriented towards improving health outcomes is a lack of data, which is exacerbated by strict controls over databases held by big pharma companies, plus the high resource and time costs of experiments and optimization processes. Furthermore, as with healthcare above, the literature highlights the implications of bias and discriminatory data and emphasises the role of ethics and interpretable AI in progressing the field.

3.2.4 Insights from the literature

This review demonstrates that applications of AI in Science and Technology can be found across all STEM fields. AI presents many potential opportunities to transform the way research is done, and to extend the capabilities of industries associated with the life sciences, health sciences, and physical sciences. The enthusiastic adoption of AI, however, should be tempered with a view to managing risks and unintended consequences associated with AI. Of particular concern are data systems and practices that discriminate against particular social groups. There are concentrations of AI use in digital health, medicine and dentistry, environment, materials science, engineering, computational and mathematical sciences, robotics, quantum information technologies, health informatics, computer aided design, nurse and health care professional education and training, sustainability, dynamical and control systems engineering, safety critical systems, genetics, agriculture, energy, conservation and efficiency, data analytics, mechanical engineering, smart innovation and computer and information science. Crucially, the application of AI in science and technology looks to consider its application from a range of perspectives, indicative of the AI ecosystem comprising actors across the research community, policy and practice.

The use of AI in certain domains, particularly to support resource allocation decisions (e.g. in health) and to monitor critical infrastructures and processes (e.g. in engineering), is especially high stakes. As such, much of the material we describe inevitably refers to ethical implications in their findings and conclusions, and engages with the broader literature on ethical implications and governance of ML.

Further, there is a clear theme relating to the collective need for multi and interdisciplinary working practices across the research landscape, and the development of shared data infrastructures. In AI, this extends further too, with research involvement from public and private sectors who comprise the AI ecosystem. While it is clear that AI can transform STEM, caution should be taken in applying a broad brush to its implementation. This is where AI futures and related work in the fields of the humanities and social sciences can support research and development, addressing issues of biases and inequalities emerging from the field of AI research.

3.2 Detailed report on literature across AI in STEM

3.2.1 AI in the physical sciences

3.2.1.1 Engineering

Use of AI in chemical engineering has a long history, which extends back to the 'expert systems' era of AI research (Venkatasubramanian 2019; Schweidtmann, 2021). This history is driven by the central importance of modelling to the chemical engineering field: models critical for understanding how chemicals react, managing chemical processing, etc (Venkatasubramanian, 2019; Dobbelaere, 2021). Notable successes in the application of AI in chemical engineering include: industrial applications of AI to process operations and diagnosis (Venkatasubramanian, 2019; Schweidtmann 2021); 'inverse design' of materials to meet desired properties (Venkatasubramanian, 2019); design and discovery of catalysts (Venkatasubramanian 2019); and, predicting quantum chemical properties (Dobbelaere, 2021). Emerging areas of research include the development of hybrid physicochemical-data-driven models for improved interpretability, extrapolation, and prediction accuracy (Schweidtmann, 2021). Adoption of AI in the chemical industry is primarily driven by two factors: a technology push, and an industry pull (Schweidtmann 2021). On the technology push side, there is a profusion of data available for training of predictive models (Dobbelaere, 2021; Schweidtmann, 2021) and a host of relatively easy-to-use open-source tools for application of ML techniques (Venkatasubramanian, 2019; Dobbelaere, 2021; Schweidtmann, 2021). Meanwhile, on the industry pull side there are high levels of industry competition, and environmental/regulatory drives towards efficiency, that create strong incentives for adoption of ML to automate and optimise chemical processing (Schweidtmann, 2021). However, challenges to further adoption of AI in chemical engineering remain. First, although chemical engineering data is increasingly available, this availability is still patchy, especially compared to other AI application domains (e.g. NLP) (Schweidtmann, 2021). Second, the representation of chemical engineering data in numerical data is an ongoing area of research (Dobbelaere, 2021; Schweidtmann, 2021). Third, although ML tools are accessible to chemical engineers, substantive training in ML techniques is yet to be integrated into chemical engineering training (Venkatasubramanian, 2019).

3.2.1.2 Chemistry

As in Chemical Engineering, the field of Chemistry has a long history of adopting AI techniques, which predates the current era of Deep Learning (Mater, 2019; Gasteiger, 2020). Indeed, the field of Chemoinformatics, which emerged in the 1960s in parallel to the profusion of chemical structure and property data, has always been closely tied to the field of AI (Gasteiger, 2020). Nonetheless, the success of Deep Learning techniques in many chemistry subfields has led to a significant increase in use of AI across the field of chemistry since 2015 (Baum, 2021). In particular, Deep Learning techniques are being used in retrosynthesis of chemicals, reaction optimisation, and drug design – all problems that traditional computing approaches were unable to advance (Mater, 2019). Indeed, Deep Learning techniques have been applied in all stages of the chemical research workflow, from modelling of molecular structures to design of molecules to chemical synthesis (Matter 2019). These techniques, and AI more generally, have been adopted in a wide range of industries, and many have developed into chemistry subfields in their own rights (Gasteiger, 2020). In particular: computer-based drug discovery is now a significant field, with close ties to the pharmaceutical industry; agricultural research makes use of AI to develop agrochemicals with desirable properties (e.g. low toxicity, with the field of chemoinformatics playing a large role; food science makes use of AI to map the relationship between chemical structures and properties of food; cosmetics makes use of AI to predict toxicity of chemicals (replacing animal testing) and to design molecules for new cosmetic products; materials science makes use of AI to design materials with specific properties; process control in industrial chemical manufacturing makes use of AI to detect faults in chemical processes; and, the field of regulatory science is increasingly using AI to predict toxicity of new chemicals (Gasteiger, 2020). The proliferation of AI in chemistry since 2015 is due to three factors: rapid growth in computing power; access to open-source Machine Learning frameworks; and increasing data literacy among chemists (Baum, 2021). These factors, combined with the field's existing data infrastructure, mean it is likely AI techniques will continue to play a significant role in chemistry research into the future (Baum, 2021).

3.2.1.3 Computer Science

Artificial intelligence is a sub-discipline of computer science. The field of computer science covers all aspects of AI, and is concerned with the development of new techniques as well as their applications. It has generated AI techniques with a wide variety of uses, including: a) machine learning, b) neural network and deep learning, data mining, c) knowledge discovery and advanced analytics, d) rule-based modelling and decision-making, e) fuzzy logic-based approach, f) knowledge representation, g) uncertainty reasoning, h) expert system modelling, i) case-based reasoning, j) text mining and natural language processing, k) visual analytics, computer vision and pattern recognition, and l) hybrid approach, searching and optimization (Sarker, 2022). The field of computer science focuses on each of the application areas described in the physical sciences, health sciences, life sciences described here.

3.2.1.4 Earth and Planetary Sciences

As fields that rely on vast quantities of data to track natural and anthropogenic phenomena, the Earth and Planetary Sciences make wide use of AI technologies (Sun, 2022; Tahmasebi, 2020). Indeed, practitioners in the Earth and Planetary Sciences are contributing to the development of new AI technologies through their identification of limitations with existing AI models (Zhong, 2021). Use of AI is widespread across all areas of geology, particularly in the search for minerals and energy, and the prediction of natural phenomena (e.g. earthquakes, wildfires, and rainfall) (Sun, 2022; Zhong, 2021). A particular area of focus is the application of AI models to satellite data (Tahmasebi, 2020). Here, advances in AI techniques for image enhancement and processing, modelling of fluid dynamics, and multi-modal models is driving the application of AI for climate pattern recognition, particulate matter identification, modelling of groundwater systems and rainfall-runoff, flood prediction, optimisation of energy extraction processes, and monitoring of carbon dioxide leakage (Tahmasebi, 2020; Zhong, 2021). Ongoing research is focused on bridging the gap between geoscientific data and existing AI models, and on the development of planetary-scale systems to monitor and forecast nature, support human society to adapt to environmental changes, and guide better decisions about natural resources (Sun, 2022). This includes adopting state of the art Deep Learning techniques, use of physics-informed AI models, improving the scientific interpretation of model outputs, and improving benchmarking of AI models used in the Earth and Planetary Sciences (Tahmasebi, 2020). Barriers to further adoption of AI in the Earth and Planetary Sciences include: interpretation of Deep Learning models, which is vital for knowing if model predictions conform to fundamental principles of the Earth Sciences; lack of mature databases for many environmental applications; and, improving scientists' understanding of AI (Zhong, 2021).

3.2.1.5 Energy

Within the All Science Journal Classification schema, the Energy subject covers research associated with energy use, management, and extraction. In this review, we consider the use of AI in renewable energy, oil and gas, sustainability, and logistics research and management. Across these subfields, AI is now widely used in modelling and forecasting. In the renewable energy field, forecasting of power generation is a major problem due to the inherent uncertainty of renewable energy sources (Wang, 2019). AI techniques have been used to address this problem in solar and wind energy (Nishant, 2020), and ongoing research is focused on improving the use of AI to predict wave, geothermal and other renewable energies (Wang, 2019). Ongoing research is also focused on the development of physics-informed models and unified predictive models, that make use of multiple forms of data (e.g. energy demand and weather data) to develop more accurate and real-time predictions of renewable energy demand and supply (Wang, 2019). In the sustainability field, AI techniques have been similarly used to improve forecasting, particularly of water resource conservation and of different climate mitigation strategies (Nishant, 2020). In the oil and gas field, AI technologies have been directed towards improving efficiency of exploration and extraction of natural resources (Koroteev, 2021; Tariq, 2021). Specific applications include: speeding up manual mapping of reservoirs or typing of rock samples and identification of human errors; detecting drilled rock type and potential drill failures through real-time drilling telemetry and analysis; simulation of reservoir development scenarios; and forecasting the efficiency of different extraction strategies (Koroteev, 2021; Tariq, 2021). In the logistics field, AI techniques have also been used to drive efficiency gains, particularly in strategic and tactical process optimisation, predictive maintenance of assets, decision support systems for asset allocation, production planning, and operational process optimisation (Woschank, 2020).

3.2.1.6 Environmental Science

AI has been enormously useful in the environmental sciences, given its ability to deal with massive volumes of environmental data from satellite and in situ sources. Numerous subfields have benefitted from these technologies, including those of biodiversity, renewable energy, water and resource conservation, sustainable transportation, smart cities and climate change (Nishant et al., 2020). There are several areas of interest across these subfields. These include work on Connected and Autonomous Electric Vehicles, which focus on self-driving capacity, advanced communications, and enhanced mobility that can reduce environmental burdens of transport, as well as conservation biology, where AI has been able to rapidly process a range of signals, identify risks, and provide real-time conservation alerts, such as risks to different wildlife populations (Gupta et al., 2021). Significant strides have been made in the science of weather prediction and climate prediction (Boukabara et al., 2020). For example, AI technologies have been applied to weather prediction, meteorology and oceanography, sunspot detection, and automated detection of extreme weather events (Boukabara et al., 2020). There has also been a push to use AI’s monitoring capacities, to monitor soil, water and air, with the aim to improve and preserve healthy environmental conditions. This field also focuses on rapid and substantial reinvention of mainstream industries and processes, in order to reduce their impacts on climate change, food and water insecurity, ecosystem degradation, and societal breakdown (Gupta et al., 2021). Key challenges for how humans respond to AI-based interventions, risks to cybersecurity, the potential for negative impacts of AI applications, and uncertainty in measuring effects of interventions (Nishant et al., 2020).

3.2.2 AI in the health sciences

The use of technology for communication and data gathering have long been used to enhance healthcare services. Its adoption in clinical use reveals disparity in communication between AI scientists and medical personnel (Rathinam et al., 2021). Notwithstanding, AI offers huge potential for improving healthcare in supporting clinical decision making and new techniques for diagnostics and detection as well as clinical evaluation. However, its deployment is often opaque, and adoption can be 'inhibited by the use of 'black box' AI systems where it is not understood why AI is an effective technique (Paton & Kobayashi, 2019) or how far it protects the rights of patients to confidentiality, consent and autonomy. Indeed, the rapid integration of AI in health has occurred with little communication between those developing it, computer scientists and doctors (Straw & Callison-Burch, 2020) where, for instance, historical biases may widen healthcare inequalities. Sapienza et al. share one such example, where in their review of urban interventions and adopting digital technologies and AI-based algorithms to improve population health (2022), they show that out of 3733 records screened only 12 papers met their inclusion criteria to assess this with only one article using a comprehensive approach to public health "investigating the use of AI and digital technologies both to characterise exposure conditions to health determinants and to monitor population health effects, while the others were limited to characterising exposure conditions to health determinants, thus employing a preliminary public health perspective." This is suggestive of a need for a more comprehensive approach to the use of technologies in sustainable living and health. This also highlights the ethical implications for both policy and practice. Such debates are out of the scope of this review but crucial critical context as ethics appears concomitantly with reviews of healthcare - ethics and healthcare are intertwined. The question remains as to whether AI can be implemented successfully in medicine and health against the structures of governance that exist within it, for patient benefit (Rathinam et al., 2021) and in balancing the trade-offs, where for some fields, AI is transforming healthcare, indeed, in biotechnology, it is argued that 'in the future, no biotechnology can do without AI' (Holzinger, 2023).

AI is used widely across the health sciences, with applications in and across key domains such as mental health, pharmaceutical industries, physiology, neuroscience, infectious diseases (notably a rise in journal articles in lieu of the Covid 19 pandemic), cardiology, genomics, migraine and chronic headache, lower back pain, forensic science, CT imagery, support for intensive care units, Chat GTP, regenerative medicine, drug screening, brain tumour research, skin disease, fitness, bioinformatics, urology, dementia, dermatology, decision making in nursing, diabetes, spinal trauma, synthetic biology, neurology, stroke research, surgery, urban health, ophthalmology, biotechnology, dentistry, cancer and stem cell research, radiography, perinatal, obstetrics and child health care as well as psychology. Most commonly these areas of application are explored through the lens of disease detection, prediction, modelling, diagnosis, treatment, discovery, robotics, decision making and health care 'futures' such as considerations of major grand challenges such as ageing population and obesity. For instance, where the use of wearables and smart homes are reviewed and shown to support physiological monitoring, emergency detection, safety monitoring, social interaction and cognitive assistance of the elderly (Sorwar & Hoque 2021).

As digital technologies and AI transform the health sector across the globe, there have been efforts to conceptualise the definition of 'digital health'. Fatehi, Samadbeik & Kazemi (2020) found that more concerted research effort has been made in showing the provision of healthcare rather than the use of technology itself where the dominant concept of digital health relates to mobile health and AI. For the purposes of this review, we include the concepts of digital health care, terms, AI, ML, NLP and DLM. The overall theme of AI in health care suggests the need for a more integrated approach, where AI tools can be used as more 'objective measures', if combined with patient reported views and outcomes and value-based healthcare (Raclin et al., 2022). For instance, mapping health care priorities through objective measures using AI and machine learning if combined with patient reported outcomes and value-based health care reinforce the need to involve the public in the development of machine learning systems in health care as well as more multi-disciplinary working to increase communication, transparency and explainability.

3.2.2.1 Medicine

The adoption of AI in medicine is wide-ranging. Whether it is developing a digital lung CT AI in clinical medicine (Newell, 2024) to the ethical use of AI in radiology, therapeutics for COVID-19 to transplantation pathology and mental health. One clear example of AI in medicine comes from the use of AI during the COVID-19 pandemic. (Norozpour, 2021) explores a review of AI and its relationship to modelling and simulation in health during COVID-19. The study has involved a critical review of different pieces of literature on the value of artificial intelligence in modelling and simulation and also finds out the model of the relationship between the COVID-19 death rate and the number of handwashing materials. AI was also used in hospital management. Khanam, describes how AI-based decision making would support managing patients with COVID-19 more efficiently by using AI methods to enhance their critical care (2022). For instance, the development of an AI based model enabled clinicians working on a vaccine, testing facilities etc could then be supported by a decision-making tool to help with diagnosis, treatment and risk stratification as well as clinical management. During COVID-19 AI was used as a prognosis tool in patients using lab tests (Khounraz, 2023). AI was used in predicting order processing times in E-pharmacy supply chains and in predicting eligibility for vaccines (Bisht et al., 2023) and has emerged as a promising tool for facilitating resource distribution, especially during medical emergencies (Wu & Wang, 2023). Further, a review of COVID-19 and AI revealed potential functionalities of such technologies that can be used to predict mortality, detect, screen, and trace current and former patients, analyse health data, prioritise high-risk patients, and better allocate hospital resources in pandemics, and generally in health-care settings, for instance. The use of AI and robotics in ophthalmology and cyber surgery revealed 68 articles where robotics and AI have been used to perform repairs in eye surgery including cataract surgery (Alafaleq, 2023) removal of melanoma. Here robotics is discussed as an alternative to human surgeons because of shortage though these technologies are in their infancy and highly expensive, scarce in availability and ethically dubious with respect to safety and precision. AI has also become an important aspect of plastic surgery. Big data, machine learning, deep learning, natural language processing, and facial recognition are examples of AI-based technology that plastic surgeons may utilise to advance their surgical practice, but a review of these applications reveals important ethical considerations around patient autonomy, consent and confidentiality (Jarvis et al., 2020).

There is a particular emphasis on the role of AI in oncology. Mysona et al., review the use of AI in gynaecologic oncology, where AI is used to advance tailored screening, precision surgery and personalised therapies (2021). There has been a significant rise in research in this area in the past 20 years where AI can be seen to enhance diagnosis, refine clinical decision making and create more personalisation. The issue of its rapid adoption, comes with the consideration of data quality, interpretation and transparency and that a better understanding of the computer science behind the algorithms, would support physicians and patients. AI is used in a wide spread way to identify cancer through imaging though caution is stressed in its use because of the reliance on human judgement - one of the core tenets of medicine (Bi et al., 2019). However, AI is seen to support detection, characterisation and monitoring of cancers. AI can automate processes in the initial interpretation of images. However, the interpretation of the large volume of data that is generated by these advancements "presents a barrage of new potential challenges". The curation of data sets and increasing transparency about how these 'black boxes' work, perhaps through data visualisation will allow a degree of explainability to those working in healthcare when understanding how algorithms make decisions.

Added to this literature review is the field of mental health and psychiatry. AI is used in mental health care where large language models are being used in psychiatry. However, its use is cautioned as it is seen to be embedding harms and biases with respect to religion, race, gender, nationality, sexuality and age (Straw & Callison-Burch, 2020). The use of language models in particular means that narratives for assessing mental health can be used to provide rich information on emotional and psychological wellbeing of patients (Conway et al., 2019). For instance, Natural Language Processing (NLP) models have been used to predict suicidal ideation, post-partum depression using online data and self-harm using social media mining. At present, researchers and developers are building these tools with the assumption that existing medical practice is 'gold standard', despite the field's long history of discriminatory practice, biases and medical error (Hamberg & Medicinska Fakulteten, 2008). Similarly, a range of medial biases can seep into the use of tools for predicting personality disorders and post traumatic diagnoses. Such papers highlight how digital professionals can prevent the exacerbation and projection of media bias in digital health. AI has been applied to the area of weight loss and obesity, a review of the potential for AI in enhancing adult weight loss shows that (AI) could be used to regulate eating and dietary behaviours, exercise behaviours and weight loss (Chew, Ang & Lau, 2021). Here, machine learning perception can focus on recognising food items, eating behaviours and physical activity. It can also predict weight loss and dietary lapses as well as emotional eating related to online nudging and personalised prompts online. Such use is generally seen as beneficial with some warning concerning its contingency on engagement and contextualisation.

3.2.2.2 Dentistry

AI has widespread use in dentistry. Applications of AI across dentistry show the benefits of this technology where these tools pay special attention to the area of aesthetic dentistry and colour research (Carrillo-Perez et al., 2022). 'Digital dentistry' provides personalised treatments for patients, and aesthetic dentistry is shown to benefit patients enhancing accuracy of dental restorations and advances in tooth colour. A further review provides a comprehensive look at evaluating the diagnostic and prognostic accuracy of artificial intelligence in endodontic dentistry (Karobari et al., 2023): "AI technologies have primarily been used in dentistry to diagnose dental diseases, plan treatment, make clinical decisions, and predict the prognosis. AI models like Convolutional Neural Networks (CNN) and Artificial Neural Networks (ANN) have been used in endodontics to study root canal system anatomy, determine working length measurements, detect periapical lesions and root fractures, predict the success of retreatment procedures, and predict the viability of dental pulp stem cells" (Karobari et al., 2023). Here, AI can support disease prediction, modelling and improves accuracy in terms of diagnostics.

3.2.2.3 Veterinary science

Recent developments in the field of machine learning (deep learning, ensemble learning, voice recognition, emotion recognition, etc.) are still relatively new in the field of veterinary (Cihan et al., 2017). The majority of this work is found mostly in the area of agriculture though it does extend to vet practice. AI can be used in decision support systems for instance, to look at reproduction and livestock farming, estimates of insemination in dairy cows, to estimate egg fertilisation in livestock, classifying infectious disease risk across animals, to automate and identify risk in swine fever, determine feeding performance of animals and estimate body weight of livestock. AI is used in assessing the welfare of livestock animals, which can improve their wellbeing. While affective computing in human research has received increasing attention, the knowledge gained on human emotions is yet to be applied to non-human animals (Neethirajan, 2021). In vet practice, AI is used in diagnosis of animal disease, disease decision support and to predict animal diagnoses. The concentration of ML and AI in farm and herd management comes with risk and that many computer scientists have not yet fully addressed the problems encountered by the veterinary field. A multidisciplinary approach is called for as the potential for AI in vet practice is seen as 'unmet' Basran & Appleby (2022). The impact of AI on vet education and practice is also cited as an area where AI could be used by the European Coordinating Committee of Veterinary Training[3]. Here, AI's potential impact in vet science is reported to be mostly found to improve communication and information with clients and stakeholders, improve cross-disciplinary working, enhance the development of prevention strategies and drug and vaccine development as well as improve prevention, diagnosis and treatment of animal and zoonotic diseases.

3.2.2.4 Healthcare professionals

The literature clearly suggests a need for more infer and multi-disciplinary dialogue across AI scientists and clinicians. (Rathinam et al., 2021) explore the key challenges that AI researchers face in terms of persuading medical practitioners to utilise AI systems for clinical practice. Industry best practices such as the development of specific quality standards and enhancing interdisciplinarity could help to bridge the gap between computer scientists, medical practitioners and medical administrators. These kinds of gaps in communication, transparency and explainability emerge as a dominant theme in the literature. AI can assist with some of this, if used in combination with patient views and outcomes, but also through its use in surgical and medical education. For instance, the use of AI in surgical education reveals 49 studies reviewing AI intervention AI in surgical education, particularly for the assessment of surgical competencies of trainees across dentistry, postgraduate training and surgical fellows (Kirubarajan et al., 2022).

3.2.3 AI in the life sciences

AI and the life sciences have benefited from strongly intertwined progress. The sections below show the use of AI in the five fields: 1) agriculture and biological sciences, 2) biochemistry, genetics and molecular

[3] Accessible at:
https://www.intranet.eaeve.org/fileadmin/downloads/eccvt/DTAI_WG_final_report_ECCVT_adopted.pdf

biology, 3) immunology and microbiology, 4) neuroscience and 5) pharmacology, toxicology and pharmaceutics.

3.2.3.1 Agricultural and Biological Sciences

Modern agriculture is tasked with producing more food while addressing climate change, natural resource depletion and health and safety concerns. Thus, research in this field plays an important part in improving effectiveness and reducing environmental burden. There are already a range of machine learning applications in agriculture. Indeed, this is an area of rapid growth, with a 164% rise in articles using AI analyses in the field between 2018 and 2020 (Benos et al., 2021). These AI-based applications include crop management (e.g. improving crop quality, matching the crop to the market, identifying diseased crops and weeds), water management (e.g. monitoring the status of soil water, crop growth conditions, weather conditions), soil management (e.g. remote sensing and soil mapping), and livestock management (e.g. monitoring the quality and living conditions of animals) (Benos et al., 2021). AI has been used in each stage of farming: pre-harvesting (e.g. improving seed quality, reducing fertiliser/pesticide application, optimising irrigation), harvesting (e.g. improving fruit classification, taste, firmness) and post-harvesting (e.g. reducing use of chemicals, improving shelf life and fruit handling processes) (Meshram et al., 2021). One key area of AI relates to the use of autonomous farming machinery and analysis of large volumes of data on present and past conditions to improve predictions of crop diseases and pest infestation (Qazi et al., 2022). In this field, the combination of hardware and software has been key. Durable, inexpensive and power-efficient hardware can be installed to measure different aspects of the farming process, while withstanding harsh climate conditions. AI and big data software are then able to analyse the large volumes of data accumulated by these hardware models, while combining insights from the latest scientific trends (Qazi et al., 2022). Together, this combination improves farmers' control, allowing for more precise use of water, fertiliser, and pesticides while optimising crops, supply chains, and sustainability (Misra et al., 2022). The use of AI in agricultural research can improve efficiency, reduce waste, and increase food security, making it a valuable tool in the effort to feed a growing global population.

3.2.3.2 Biochemistry, Genetics and Molecular Biology

AI is now frequently used for classification and prediction of various biochemical, genomic and molecular biological processes and phenomena. Biochemical research has benefited from AI-assisted analysis of protein structure and function. AI algorithms can predict the 3D structure of proteins based on their amino acid sequence, which can help researchers understand how they interact with other molecules in the body (Callaway, 2020). In the area of clinical biochemistry, AI has been used to improve laboratory processes and care pathways (enabling prediction of scenarios and optimization of task management), imaging and laboratory services (assisting ordering tests based on diagnosis and triggering alerts for abnormal results) and diagnosis and disease monitoring (providing decision support via integrating patients' histories, clinical records, ongoing interventions and imaging/laboratory results) (Gruson et al., 2019). Similarly, advances in AI have been seen in the field of genetics. For example, AI-assisted genetic sequencing of tumours has enabled the personalised treatment of cancer (Luo et al., 2020). AI applications in clinical genomics have sought to improve tasks that are impractical and error-prone when using standard statistical approaches. Many of the AI techniques relate to various steps involved in genomic analysis. These include variant calling (identifying individual genetic variants among the millions in each genome), genome annotation (identifying relevant genetic variants), variant classification (classifying potential candidate variants and analysis of non-coding sequence data), and phenotype-to-genotype correspondence (identifying pathogenic variants and determining correspondence between the diseased individual's actual phenotype and those expected to result from the pathogenic variants) (Dias & Torkamani, 2019). The goal is to move to genotype-to-phenotype predictions through AI technologies. AI models can simulate the behaviour of biological systems, such as protein folding, drug interactions, and metabolic pathways (Helmy et al., 2020). Thus, AI is increasingly important in biochemistry, genetics, and molecular biology research, helping scientists to analyse large amounts of data and make new discoveries that could lead to better treatments for diseases.

3.2.3.3 Immunology and Microbiology

AI is being used in many fields that focus on the immune system. These tools have been used in analysis, detection of relevant inputs and predicting prognosis and treatment outcomes (Jabbari & Rezaei, 2019). AI technologies have been used in detection and classification of phenotypes (individual traits) to determine the presence of a particular disease or its outcome. ML can classify virus strains upon phenotypic features, which has traditionally been a complex and tedious task when done with conventional statistical analyses. ML has also served vaccine development by helping design optimum components of vaccines (Jabbari & Rezaei, 2019). Beyond these applications at the molecular level, AI can be used at the clinical level. It can assess patient eligibility for clinical trials using electronic medical records, and can improve the likelihood of probability of enrolment in a suitable clinical trial. Similarly, deep learning has become a powerful tool to address significant challenges in microbiology. Traditionally, the field has relied on time consuming and repetitive image segmentation and classification of biomedical images. However, many imaging tasks have been transformed by AI techniques, including identifying subcellular features, enabling restoration of high-quality images from noisy data, and allowing specific cellular labels from unlabelled specimens (Chamier et al., 2019). These types of automated high-performance big-data analysis are greatly increasing the capacity for scientific study of microorganisms. Challenges for the field include the substantial resources required for generation and curation of datasets used to train the AI models, and an accessibility barrier for those without adequate resources (Chamier et al., 2019). While progress is expected to be made on this front, given increasing availability of hardware and software, there is potential for in-built biases that occur via curation of training data.

3.2.3.4 Neuroscience

The simultaneous progress in AI and neuroscience has led to a two-way relationship between the two fields (Macpherson et al., 2021). On the one hand, AI is transforming our ability to observe and manipulate brains at a large scale and to quantify complex behaviours (Richards et al., 2019). It is revolutionising comprehension of brain functions and has enabled the creation of new neural networks based on the architecture of the brain. In particular, deep learning techniques have been utilised to replicate how the cerebral cortex of the brain controls vital functions like memory, visual processing, and motor control (Kriegeskorte & Douglas, 2018). On the other, advancements in AI, particularly machine learning and neural networks, have been used to improve automated analysis of big data in neuroscience research. AI techniques are now increasingly used in the analysis of animal behaviour (e.g. automated identification of animal grooming, freezing, and social behaviour, prediction of animal behaviour based on neural activity data), in processing and classifying large image datasets (e.g. neuroimaging, histopathological images, brain tumour MRIs), analysis of brain signals (e.g. EEG signals), as well diagnosis, prediction and treatment for psychiatric, neural and developmental disorders (e.g. Parkinson's disease, epilepsy, multiple sclerosis) (Kellmeyer, 2019; Macpherson et al., 2021). Data-driven research in neuroscience stands to benefit from the growing availability of personal data and machine learning techniques. However, the increasing sophistication and autonomy of AI tools bring with them substantial ethical concerns.

3.2.3.5 Pharmacology, Toxicology and Pharmaceutics

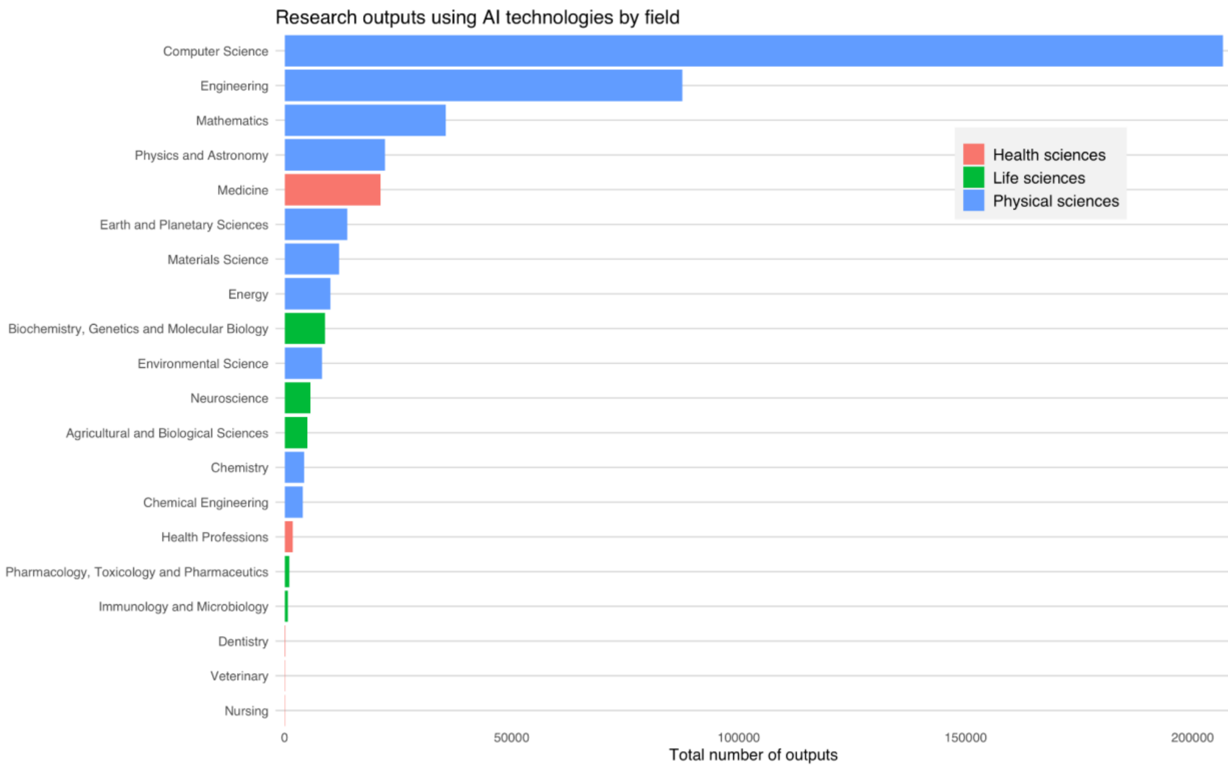
In recent years, the use of AI in the fields of pharmacology, toxicity and pharmaceutics has increased substantially. AI models, such as unsupervised clustering of drugs or patients and supervised machine learning approaches, have been useful across the spectrum of research and clinical practice. These multidisciplinary fields feed into every clinical discipline of medicine. The pharmaceutical industry has traditionally relied on trial-and-error experiments for drug formulation and delivery, which are expensive, slow and unpredictable (Wang et al., 2021). However, AI tools are increasingly used to create predictive models and recognize complex patterns in big data sets in drug research and development (Kocić et al., 2022). The use of artificial intelligence and machine learning algorithms, molecular modelling, mathematical modelling, process simulation, and physiologically based pharmacokinetic modelling has been termed "Pharma 4.0" (Wang et al., 2021). In particular, deep learning algorithms in drug design and development have taken a dominant place in the field because of improved feasibility of dealing with enormous amounts of chemical data (Kocić et al., 2022). AI technologies have been used in each stage of pharmacology. This includes drug discovery (e.g. absorption, toxicity, binding efficiency), preclinical development (prediction of drug characteristics), clinical development (participant selection and follow up), real world use (data mining) and optimising treatment (Lee & Swen, 2023). Two key areas receiving increasing attention include application of AI methods to improve the safety of pharmacotherapy and to recommend and personalise treatment. The related field of toxicology, which seeks to understand the link between drug and its final effect, is also moving to AI techniques to manage large data sets, such as mass identification and analysis of substances (Wille & Elliott, 2020). Key challenges include the potential trade-offs between interpretability and accuracy, and developing strategies for managing limited and mixed data.

4 Trends in AI publications across the STEM fields

4.1 Which STEM fields are using AI technologies?

Here, the number of articles and conference proceedings published in STEM categories 2018-2022 are shown. The data was obtained from Scopus, using fractionalised ASJC categories.

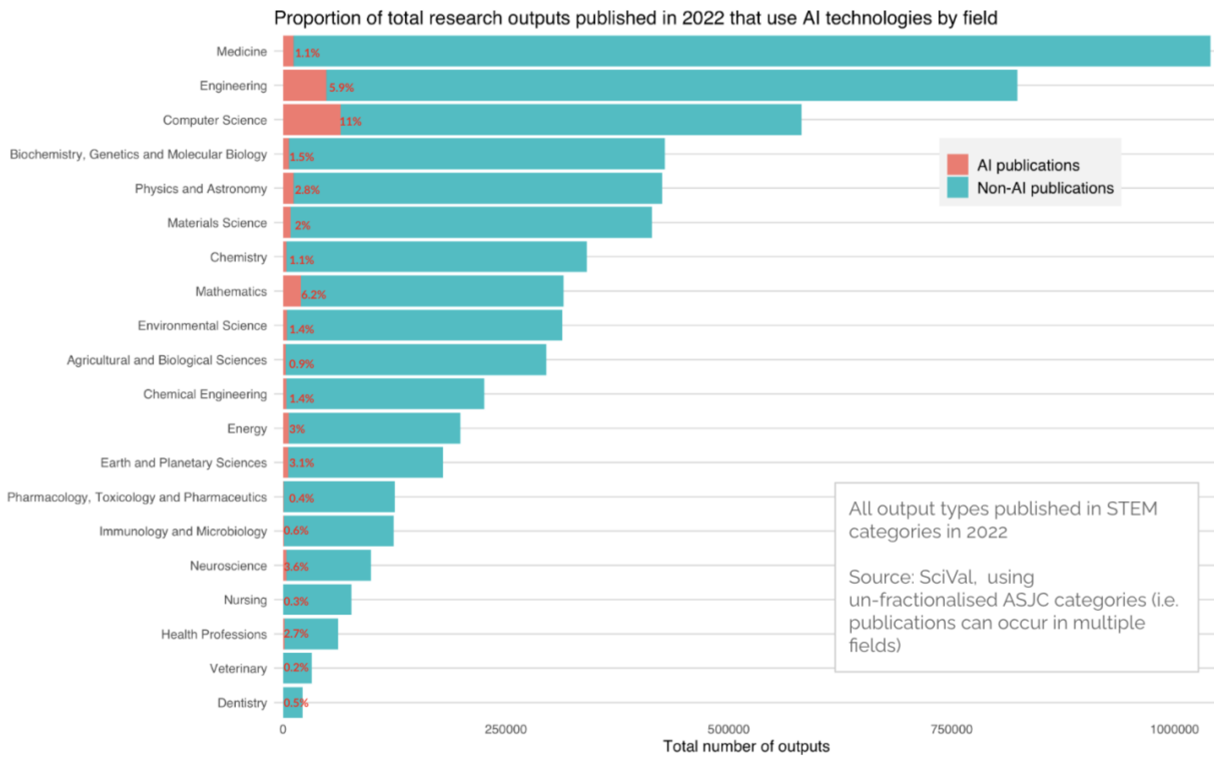
Figure 1 Outputs across fields



4.2 What proportion of fields are using AI?

Here, all output types published in STEM categories in 2022 are shown. The data is derived from SciVal, using un-fractionalised ASJC categories.

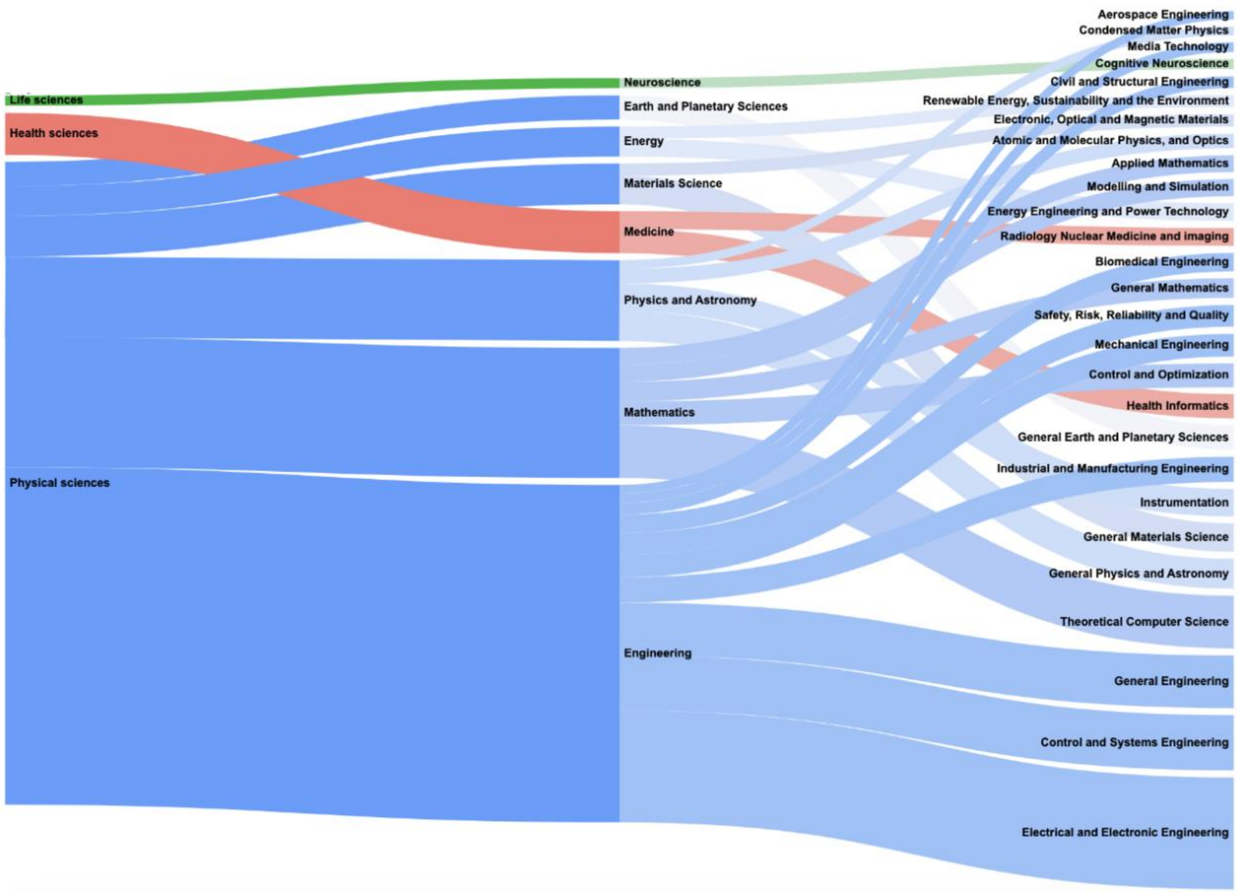
Figure 2 Research outputs in 2022



4.3 What are the main subfields using AI?

Here, subfields with over 2000 outputs are shown, based on the number of articles and conference proceedings published in STEM categories 2018-2022. The data was obtained from Scopus, using fractionalised ASJC categories.

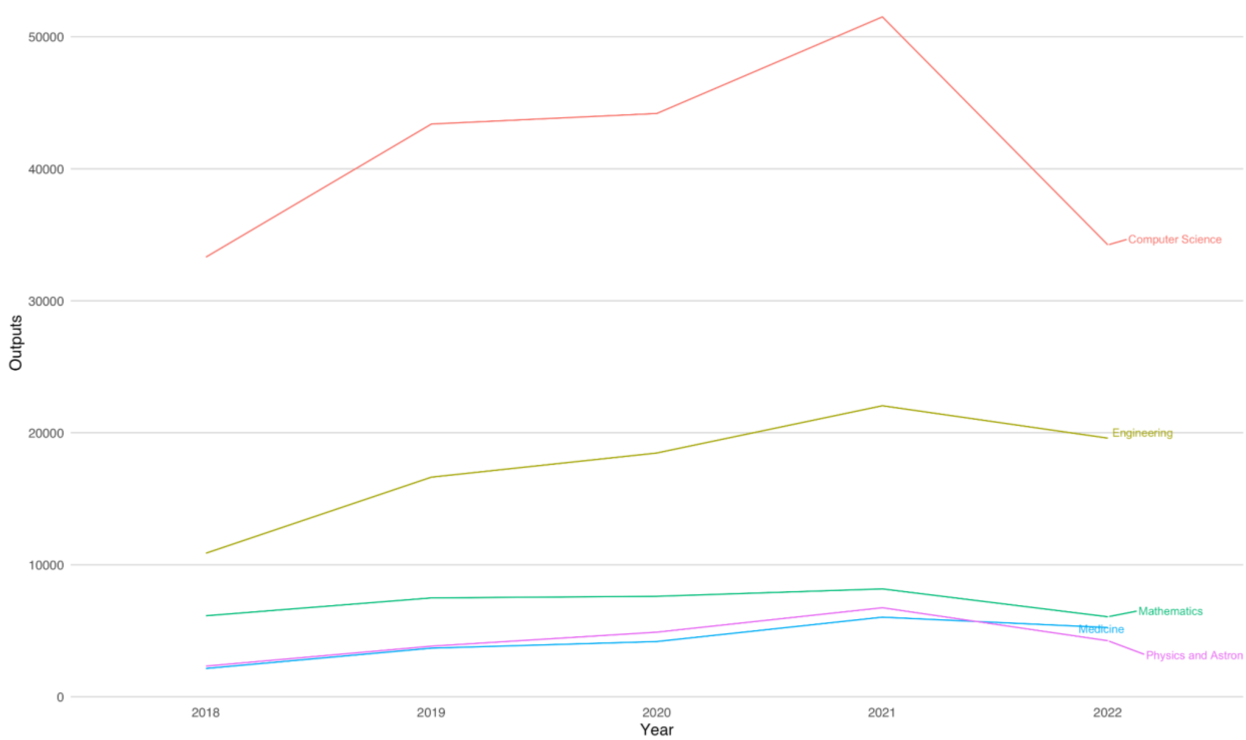
Figure 3 A Sankey diagram of outputs across fields



4.4 How have the fields using AI technologies changed over time?

Here we show the top five fields based on articles and conference proceedings published in STEM categories 2018-2022. The data was gained from Scopus, using fractionalised ASJC categories.

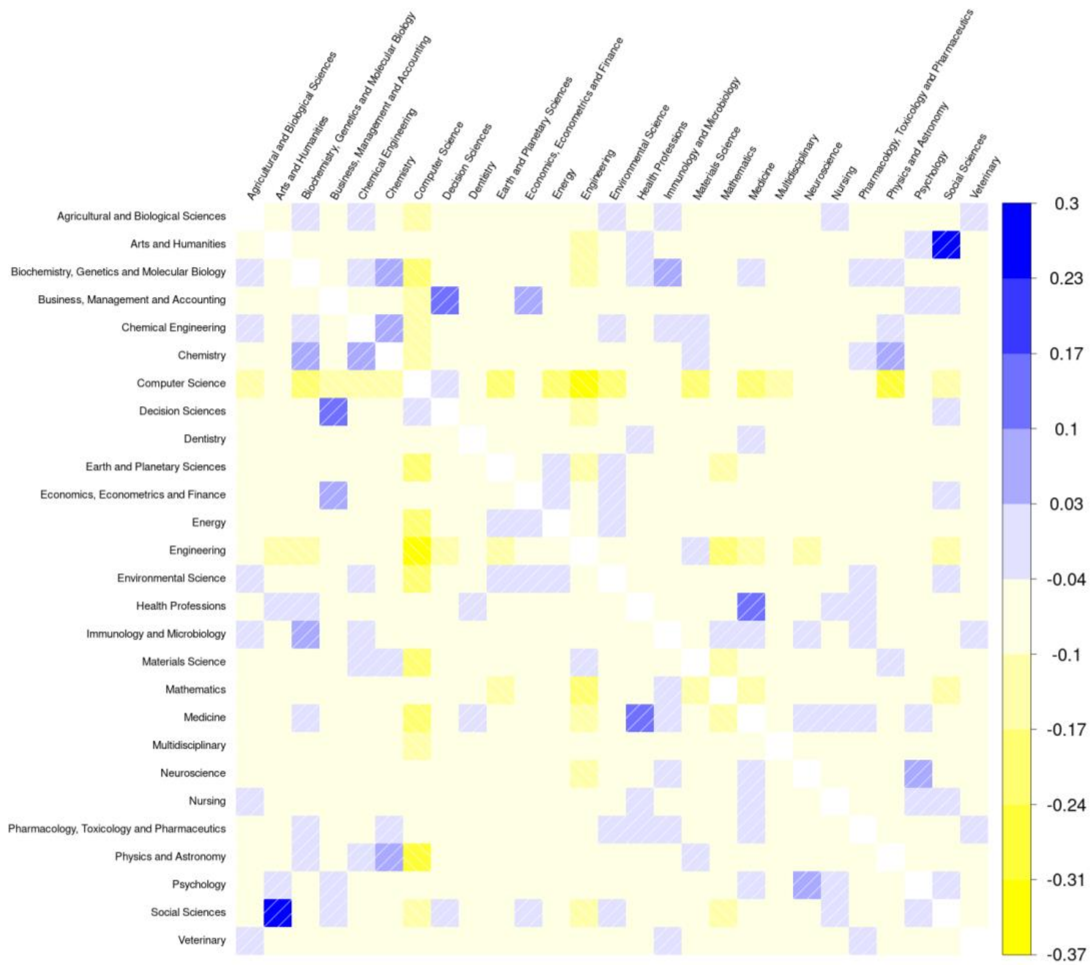
Figure 4 Top fields based on outputs



4.5 How are different fields correlated with each other?

Here we show correlations between fields (i.e. likelihood of a publication being classified under both fields) within the corpus of AI publications. Positive numbers (indicated by dark blue) represent a positive correlation, while negative numbers (indicated by bright yellow) represent a negative correlation. In general, most fields have little relationship to each other (indicated by the pale colour in most squares). This may indicate that discourse on AI is siloed within most fields. The Social Sciences and Arts and Humanities, Decision Sciences and Business, Management and Accounting, and Medicine and Health Professionals are strongly correlated. The Computer Science field is negatively correlated with most fields, indicating that AI publications in the Computer Science field are unlikely to be interdisciplinary (at least, in terms of how their fields are recorded). Further research is needed to validate whether these findings are reflective of quirks in the ASJC classification schema, or of the substantive inter-field research relationships.

Figure 5 Examining interdisciplinarity



4.6 What technologies are prominent within fields?

In this section, we show the top three AI 'topics' that occur for the five biggest fields (excluding computer science, which is used below as a comparator). A topic, taken from Elsevier's methodology, is a "dynamic collection of documents with a common focused intellectual interest" A publication can belong to only one topic.

Table 3 Prominent technologies in different fields

Fields (proportion of category)	**Key subfields (proportion of field)**	**Topics (2022 only)**
Engineering (44%)	Electrical and Electronic Engineering (31%) Control and Systems Engineering (15%) General Engineering (14%)	1. Object Detection, Deep Learning, Intersection Over Union 2. Rolling Bearing, Rotating Machinery, Failure Analysis 3. Useful Life, Health Care Management, Rolling Bearing
Mathematics (18%)	Theoretical Computer Science (36%) Control and Optimization (16%) General Mathematics (13%)	1. Object Detection, Deep Learning, Intersection Over Union 2. Generalisation, Gaussian Processes, Deep Learning 3. Active Set, Partial Differential Equation, Coefficient Inequalities
Physics and Astronomy (11%)	General Physics and Astronomy (32%) Instrumentation (30%) Atomic and Molecular Physics, and Optics (16%)	1. Object Detection, Deep Learning, Intersection Over Union 2. Rolling Bearing, Rotating Machinery, Failure Analysis 3. Hamiltonian, Ising, Qubits
Medicine (90%)	Health Informatics (27%) Radiology Nuclear Medicine and imaging (20%) General Medicine (8%)	1. Object Detection, Deep Learning, Intersection Over Union 2. Radiological Findings, Clinical Features, COVID-19 3. Texture Analysis, Cancer, Fluorodeoxyglucose
Earth and Planetary Sciences (7%)	General Earth and Planetary Sciences (42%) Geotechnical Engineering and Engineering Geology (13%) Atmospheric Science (8%)	1. Object Detection, Deep Learning, Intersection Over Union 2. Hyperspectral Imagery, Spectroscopy, Image Classification 3. Remote Sensing Image, Convolutional Neural Network, Image Retrieval
Computer Science (if treated as a field)		1. Object Detection, Deep Learning, Intersection Over Union 2. Embedding, Named Entity Recognition, Entailment 3. Collaborative Filtering, Recommender Systems, Factorization

4.7 What are the characteristics of the AI research area?

The following graphs show the general characteristics of the AI research area as a whole, based on SciVal data (2017-2022).

Figure 6 Collaboration between authors across institutions and nations

Figure 7 Collaboration between corporate- and academic-affiliated authors

Figure 8 Most prolific institutions

1.	Chinese Academy of Sciences	11,135
2.	CNRS	6,966
3.	Tsinghua University	5,913
4.	University of Chinese Academy of Sciences	5,912
5.	Anna University	5,872
6.	Shanghai Jiao Tong University	4,615
7.	Zhejiang University	4,316
8.	University of Electronic Science and Technology of China	3,759
9.	Peking University	3,463
10.	Harbin Institute of Technology	3,396

5 Taxonomy

In this section we present a taxonomy of the key AI technologies used within each field. To identify AI techniques, we used a bottom up, grounded approach and conducted a coding exercise from which to generate techniques largely inductively. From this, we state the main techniques described in the literature review and UK REF ICS by field.

The taxonomy demonstrates the prevalence of particular AI techniques across all STEM fields. Artificial Neural Networks (ANN) are mentioned in almost all fields (73%), followed by Deep Learning (63%), Internet of Things (63%) and Machine Learning (63%). Under 50% of all fields techniques referenced computer vision, Convolutional Neural Networks (CNN) (26%), robotics (26% - with a concentration in health) and big data analytics (21%).

Table 4 Taxonomy of fields and AI technologies

Fields (proportion of category)	**Key subfields (proportion of field)**	**Key technologies**
Engineering (44%)	Electrical and Electronic Engineering (31%) Control and Systems Engineering (15%) General Engineering (14%)	Machine Learning (ML), image recognition, scene understanding, decision-support systems, human centred AI, computer vision, Dynamic Linear Models, Computer Aided Design, Knowledge based engineering, Natural Language Processing (NLP), Internet of Things (IoT)
Mathematics (18%)	Theoretical Computer Science (36%) Control and Optimization (16%) General Mathematics (13%)	ML, Artificial Neural Networks (ANN), Deep Learning (DL), data mining, knowledge discovery and advanced analytics, rule-based modelling and decision-making, fuzzy logic, knowledge representation, uncertainty reasoning, expert system modelling, case-based reasoning, text mining and NLP, visual analytics, computer vision and pattern recognition, hybrid approaches
Physics and Astronomy (11%)	General Physics and Astronomy (32%) Instrumentation (30%) Atomic and Molecular Physics, and Optics (16%)	ML, big data analytics, Support Vector Machines (SVMs), Random Forests (RFs), Graphics Processing Units (GPUs), computer vision, Deep Neural Networks, Convolutional Neural Networks (CNN), ANN, genetic algorithms, Self-Organising Maps, recurrent neural networks, autoencoders
Engineering (44%)	Electrical and Electronic Engineering (31%) Control and Systems Engineering (15%)	Machine Learning (ML), image recognition, scene understanding, decision-support systems, human centred AI, computer vision, Dynamic Linear Models, Computer Aided Design, Knowledge

Fields (proportion of category)	**Key subfields (proportion of field)**	**Key technologies**
	General Engineering (14%)	based engineering, Natural Language Processing (NLP), Internet of Things (IoT)
Medicine (90%)	Health Informatics (27%) Radiology Nuclear Medicine and imaging (20%) General Medicine (8%)	DL, ML, robotics, large language models, social media, datamining, text mining, predictive analytics, classification, data analytics, IoT, chatbots, genetic algorithms, image processing, clinical decision support systems, descriptive analytics, e-health, healthcare informatics, prescriptive analytics, NLP, Chatbots, AI-based radiology reporting, COVID-19 tracking
Earth and Planetary Sciences (7%)	General Earth and Planetary Sciences (42%) Geotechnical Engineering and Engineering Geology (13%) Atmospheric Science (8%)	DL, 'physics-informed' models, remote sensors, ML, deep neural networks, computer vision, Local Interpretable Model-agnostic Explanations, image recognition, object recognition, ANN, CNN , multi-modal with social media, interactive and semantic ML, NLP
Materials Science (6%)	General Materials Science (57%) Electronic, Optical and Magnetic Materials (25%) Biomaterials (7%)	ML, big data analytics, genetic algorithms, feature generation, Bayesian frameworks, statistical models, ANN, functional recognition imaging, DL, GPU, pattern recognition, CNN
Energy (5%)	Energy Engineering and Power Technology (42%) Renewable Energy, Sustainability and the Environment (29%) Fuel Technology (11%)	ML, robotics, smart grids, VR, big data analytics, DL, computational intelligence (CI), soft computing and distributed artificial intelligence, fuzzy systems, intelligent agents, blockchain, IoT, robotics, energy optimisation, computational efficiency, predictive maintenance
Biochemistry, Genetics and Molecular Biology (42%)	Biochemistry (19%) General Biochemistry, Genetics and Molecular Biology (14%) Internet of Things (IoT) Technology (12%)	ANN, decision trees, regression analysis, biosensors, big data, NLP, data mining, biomedical ontologies, knowledge-based reasoning, decision support, spatial reasoning, DL, symbolic AI
Environmental Science (4%)	Water Science and Technology (18%) General Environmental Science (17%) Environmental Engineering (13%)	Big data analytics, ML, IoT, smart sensing technology, ANN, fuzzy inference systems, genetic algorithms, image recognition
Neuroscience (27%)	Cognitive Neuroscience (43%) General Neuroscience (32%) Neurology (7%)	Neural computation, DL, ANN, reinforcement learning, VR, robotics, temporal-difference (TD) methods, signal and image processing, CNN , neuroimaging techniques "elastic" weight consolidation (EWC), object recognition

Fields (proportion of category)	**Key subfields (proportion of field)**	**Key technologies**
Agricultural and Biological Sciences (24%)	Agronomy and Crop Science (17%) General Agricultural and Biological Sciences (16%) Food Science (13%)	IoT, fuzzy logic, ANN, management-oriented modelling, digital elevation modelling, characterisation, high order neural networks
Chemistry (2%)	General Chemistry (32%) Analytical Chemistry (31%) Physical and Theoretical Chemistry (14%)	IoT, biosensors, image recognition, mass spectrometry, sensors, chromatography, imaging, Near Infrared Reflectance (NIR)
Chemical Engineering (2%)	General Chemical Engineering (34%) Fluid Flow and Transfer Processes (23%) Process Chemistry and Technology (19%)	Training of predictive models, Open Source ML techniques, NLP, computer-aided synthesis, surrogate modelling, objective optimisation algorithms, ANN, expert systems, genetic algorithms, fuzzy logic
Health Professions (77%)	Health Information Management (43%) Radiological and Ultrasound Technology (42%) Physical Therapy, Sports Therapy and Rehabilitation (5%)	DL, image recognition, caption generation, and speech recognition, representation learning, ML, representation learning, segmentation, ANN
Pharmacology, Toxicology and Pharmaceutics (5%)	Pharmacology (24%) Drug Discovery (22%) Pharmaceutical Science (21%)	NLP, unsupervised clustering of drugs, supervised ML for drug monitoring, automation, ANN, digital twins
Immunology and Microbiology (3%)	General Immunology and Microbiology (48%) Immunology (17%) Microbiology (16%)	SVM, computer vision, supervised ML, image recognition, DL, bacteriology automation, digital storage, slide scanning
Nursing (1%)	General Nursing (25%) Leadership and Management (20%) Nutrition and Dietetics (19%)	Robotics, speech recognition, NLP, DL, mobile health applications, sensor-based technology, tele-nursing, Chat GTP, image recognition, machine learning, algorithms (social media)
Dentistry (1%)	General Dentistry (64%) Orthodontics (16%) Oral Surgery (12%)	CNN, ANN, VR, robotics, computer vision, object detection
Veterinary (1%)	General Veterinary (83%) Food Animals (10%) Small Animals (2%)	DL, ensemble learning, voice recognition, emotion recognition, image recognition, 3D imaging, automated analysis of x-ray radiographs, supervised learning, SVMs, decision trees, naïve Bayes, logistic regression, and linear regression, ANN, NLP

6 Conclusion

This report demonstrates that applications of AI can be found across all STEM fields posing many potential opportunities to transform and optimise performance of sectors. Such use should be cautioned with a view to managing risk and unintended consequences which could be deleterious to society and individual groups through discriminatory data systems and practices.

There is a concentration in digital health, medicine and dentistry, environment, materials science, engineering, computational and mathematical sciences, robotics, quantum information technologies, health informatics, computer aided design, nurse and health care professional education and training, sustainability, dynamical and control systems engineering, safety critical systems, genetics, agriculture, energy, conservation and efficiency, data analytics, mechanical engineering, smart innovation and computer and information science.

The most prominent AI techniques across STEM include Artificial Neural Networks, Deep Learning, Machine Learning, Natural Language Processing and image recognition. Other techniques that are featured regularly include Convolutional Neural Networks, Internet of Things, computer vision, robotics and big data analytics. With a growth in NLP an interesting line of future research could explore empirical accounts of researchers working in this area.

The use of AI in certain domains can be viewed as particularly high stakes, as such several of the material we describe inevitably refers to ethical implications in their findings and conclusions. Further, there is a clear theme relating to the collective need for multi and interdisciplinary working practices across the research landscape. In AI, this extends further too, with research involvement from public and private sectors who comprise the AI ecosystem. While it is clear that AI can transform these areas, caution should be taken in applying a broad brush to its implementation, this is where AI futures and related work in the fields of the humanities and social sciences can support research and development, addressing issues of biases and inequalities emerging from the field of AI research.

7 About the authors

Dr Jennifer Chubb[4]

Jennifer is a Lecturer at the University of York in the Department of Sociology. Jenn's research focuses on the role of responsibility in science and the interplay between science and society with a specialisation in Artificial Intelligence (AI). She recently published works on the role of AI in scientific research. Jenn's background is in Philosophy (University of Leeds) and in Social Science (University of York). Jenn's research expertise is in research impact, ethics, socio-technical futures, the governance of innovation, algorithms, knowledge systems, metrics, scientific narratives, and the public perception of science and technology. In recent years her research has focused on the role of responsible storytelling and ethical development of AI in the creative industries. This interest extends across a range of domains, with further specialisms in higher education and health. Jenn has published in AI & Society, Studies in Higher Education, British Politics, Palgrave Communications, The Journal of Empirical Research on Human Research Ethics and The Journal of Theory and Research in Education. Jenn completed her PhD in 2017 at the University of York where she explored the notion of instrumentalism and epistemic responsibility in science and research. She re-joined the University of York in 2019 as a research associate at the Digital Creativity Labs where she worked on AI Futures. Prior to this she was a postdoctoral research associate at the University of Sheffield, focusing on institutions, research policy, expertise in science policy, advice and diplomacy. In 2020, she began a fellowship with XR Stories at the University of York primarily focused on ethical and responsible storytelling of AI and music. Jenn is a board member of the Science and Technology Studies Unit, an executive member of AsSIST-UK and an editor for Springer Nature's Journal of Humanities and Social Sciences Communications. Jenn is an appointed advisor to the Better Images of AI project, tackling stereotypes in AI imagery and leads a network at the University of York on AI and Society, as well as a group on Algorithms, Loss and Grief. Jenn's methods are predominantly qualitative.

Dr Kate Williams

Kate is Senior Lecturer in Public Policy in the School of Social and Political Sciences at the University of Melbourne and a Visiting Research Fellow at King's College London's Policy Institute. Her research occurs at the intersection of public policy, sociology and research policy. It focuses on the production, use and evaluation of knowledge, and on cultures of evaluation and emerging methods of research impact assessment. Kate is currently lead investigator on a UK Economic and Social Research Council Research Grant (2021-2024), which compares methods and cultures of research impact in the field of AI, across the

---

[4] jennifer.chubb@york.ac.uk, corresponding author

UK, Australia, and the US. Her previous work, funded by an ESRC Future Research Leaders Fellowship and a British Academy and Leverhulme Grant, investigated the creation and assessment of value in policy knowledge contexts, including the World Bank, London School of Economics and J-PAL. She has published on these topics in the journals of Policy and Society, Research Evaluation, Policy Studies Journal and Public Administration, Big Data and Society, the Sociological Review and others. Prior to her work as Research Fellow at King's College London and consultant, Kate was a Lecturer in Public Policy at the University of York from 2019-2020. Previously, she was an ESRC Research Fellow from 2016-2019 in the Department of Sociology at the University of Cambridge and a Research Fellow at Lucy Cavendish College, Cambridge. In 2018/2019 she was a Postdoctoral Fellow at the Weatherhead Centre for International Affairs, Harvard University. Kate holds a PhD in Sociology from the University of Cambridge, funded by a Commonwealth Scholarship and the Cambridge Overseas Trust.

Glen Berman

Glen Berman is a PhD candidate in the School of Engineering and the Humanising Machine Intelligence project at the Australian National University. His research investigates the engineering tools used to research, develop, and deploy Machine Learning technologies in industry and research settings, and the relationship between tools to support Machine Learning and the practice of doing Machine Learning. Glen is currently a student researcher at Google Research, where he studies fairness interventions in Machine Learning practice, and co-convenor of the Critical Technology Studies Graduate Network, a new network of postgraduate social science researchers across Australia and New Zealand. He is the recipient of a University of Melbourne School of Social and Political Sciences Research Incubator Grant, through which he is undertaking a field-level quantitative analysis of trends in knowledge production in AI research.

8 References

Alafaleq, M. (2023). Robotics and cybersurgery in ophthalmology: a current perspective. Journal of Robotic Surgery, 1-12.

Benos, L., Tagarakis, A.C., Dolias, G., Berruto, R., Kateris, D. and Bochtis, D. (2021), “Machine Learning in Agriculture: A Comprehensive Updated Review”, Sensors, Vol. 21 No. 11, p. 3758, doi: 10.3390/s21113758.

Bisht, P., Bora, V., Poornima, S., & Pushpalatha, M. (2023). Prediction of Eligibility for Covid-19 Vaccine Using SMLT Technique. Advances in Science and Technology, 124, 462-468.

Boukabara, S.-A., Krasnopolsky, V., Penny, S.G., Stewart, J.Q., McGovern, A., Hall, D., Hoeve, J.E.T., et al. (2020), “Outlook for Exploiting Artificial Intelligence in the Earth and Environmental Sciences”, Bulletin of the American Meteorological Society, Vol. 102 No. 5, pp. 1–53, doi: 10.1175/bams-d-20-0031.1.

Callaway, E. (2020), “‘It will change everything’: DeepMind’s AI makes gigantic leap in solving protein structures”, Nature, Vol. 588 No. 7837, pp. 203–204, doi: 10.1038/d41586-020-03348-4.
Chamier, L. von, Laine, R.F. and Henriques, R. (2019), “Artificial intelligence for microscopy: what you should know”, Biochemical Society Transactions, Vol. 47 No. 4, pp. 1029–1040, doi: 10.1042/bst20180391.

Carrillo-Perez, F., Pecho, O. E., Morales, J. C., Paravina, R. D., Della Bona, A., Ghinea, R., ... & Herrera, L. J. (2022). Applications of artificial intelligence in dentistry: A comprehensive review. Journal of Esthetic and Restorative Dentistry, 34(1), 259-280.

Chew, H. S. J., Ang, W. H. D., & Lau, Y. (2021). The potential of artificial intelligence in enhancing adult weight loss: a scoping review. Public Health Nutrition, 24(8), 1993-2020.

Chubb, J., Cowling, P., & Reed, D. (2022). Speeding up to keep up: exploring the use of AI in the research process. AI & society, 37(4), 1439-1457.

Conway, M., Hu, M., & Chapman, W. W. (2019). Recent advances in using natural language processing to address public health research questions using social media and consumer generated data. Yearbook of medical informatics, 28(01), 208-217.

de Podestá Gaspar, R., Bonacin, R., & Gonçalves, V. P. (2018). Designing IoT solutions for elderly home care: a systematic study of participatory design, personas and semiotics. In Universal Access in Human-Computer Interaction. Virtual, Augmented, and Intelligent Environments: 12th International Conference, UAHCI 2018, Held as Part of HCI International 2018, Las Vegas, NV, USA, July 15-20, 2018, Proceedings, Part II 12 (pp. 226-245). Springer International Publishing.

Dias, R. and Torkamani, A. (2019), “Artificial intelligence in clinical and genomic diagnostics”, Genome Medicine, Vol. 11 No. 1, p. 70, doi: 10.1186/s13073-019-0689-8.

Dignum, V. (2019). Responsible artificial intelligence: how to develop and use AI in a responsible way (p. 59). Cham: Springer.

Gruson, D., Helleputte, T., Rousseau, P. and Gruson, D. (2019), "Data science, artificial intelligence, and machine learning: Opportunities for laboratory medicine and the value of positive regulation", Clinical Biochemistry, Vol. 69, pp. 1–7, doi: 10.1016/j.clinbiochem.2019.04.013.

Gupta, P.K., Saxena, A., Dattaprakash, B., Sheriff, R.S., Chaudhari, S.H., Ullanat, V. and Chayapathy, V. (2021), "Applications of Artificial Intelligence in Environmental Science", Artificial Intelligence: Recent Trends and Applications, CRC Press.

Hagendorff, T. (2020). The ethics of AI ethics: An evaluation of guidelines. Minds and machines, 30(1), 99-120.

Helmy, M., Smith, D. and Selvarajoo, K. (2020), "Systems biology approaches integrated with artificial intelligence for optimized metabolic engineering", Metabolic Engineering Communications, Vol. 11, p. e00149, doi: 10.1016/j.mec.2020.e00149.

Holzinger, A., Keiblinger, K., Holub, P., Zatloukal, K., & Müller, H. (2023). AI for life: Trends in artificial intelligence for biotechnology. New Biotechnology, 74, 16-24.

Jabbari, P. and Rezaei, N. (2019), "Artificial intelligence and immunotherapy", Expert Review of Clinical Immunology, Vol. 15 No. 7, pp. 689–691, doi: 10.1080/1744666x.2019.1623670.

Jarvis, T., Thornburg, D., Rebecca, A. M., & Teven, C. M. (2020). Artificial intelligence in plastic surgery: current applications, future directions, and ethical implications. Plastic and Reconstructive Surgery Global Open, 8(10).

Karobari, M. I., Adil, A. H., Basheer, S. N., Murugesan, S., Savadamoorthi, K. S., Mustafa, M., ... & Almokhatieb, A. A. (2023). Evaluation of the Diagnostic and Prognostic Accuracy of Artificial Intelligence in Endodontic Dentistry: A Comprehensive Review of Literature. Computational and mathematical methods in medicine, 2023.

Kellmeyer, P. (2019), "Artificial Intelligence in Basic and Clinical Neuroscience: Opportunities and Ethical Challenges", Neuroforum, Vol. 25 No. 4, pp. 241–250, doi: 10.1515/nf-2019-0018.

Kocić, I., Kocić, M., Rusiecka, I., Kocić, A. and Kocić, E. (2022), "Deep learning in pharmacology: opportunities and threats", European Journal of Translational and Clinical Medicine, Vol. 5 No. 2, pp. 88–94, doi: 10.31373/ejtcm/149217.

Khanam, R. (2022). Role of Artificial Intelligence in the Era of COVID-19 to Improve Hospital Management. In Global Healthcare Disasters (pp. 37-48). Apple Academic Press.

Khounraz, F., Khodadoost, M., Gholamzadeh, S., Pourhamidi, R., Baniasadi, T., Jafarbigloo, A., ... & Ayyoubzadeh, S. M. (2023). Prognosis of COVID-19 patients using lab tests: A data mining approach. Health Science Reports, 6(1), e104.

Kriegeskorte, N. and Douglas, P.K. (2018), "Cognitive computational neuroscience", Nature Neuroscience, Vol. 21 No. 9, pp. 1148–1160, doi: 10.1038/s41593-018-0210-5.

Lee, M. and Swen, J.J. (2023), "Artificial intelligence in pharmacology research and practice", Clinical and Translational Science, Vol. 16 No. 1, pp. 31–36, doi: 10.1111/cts.13431.

Luo, S., Xu, J., Jiang, Z., Liu, L., Wu, Q., Leung, E.L.-H. and Leung, A.P. (2020), "Artificial intelligence-based collaborative filtering method with ensemble learning for personalized lung cancer medicine without genetic sequencing", Pharmacological Research, Vol. 160, p. 105037, doi: 10.1016/j.phrs.2020.105037.

Macpherson, T., Churchland, A., Sejnowski, T., DiCarlo, J., Kamitani, Y., Takahashi, H. and Hikida, T. (2021), "Natural and Artificial Intelligence: A brief introduction to the interplay between AI and neuroscience research", Neural Networks, Vol. 144, pp. 603–613, doi: 10.1016/j.neunet.2021.09.018.

Meshram, V., Patil, K., Meshram, V., Hanchate, D. and Ramkteke, S.D. (2021), "Machine learning in agriculture domain: A state-of-art survey", Artificial Intelligence in the Life Sciences, Vol. 1, p. 100010, doi: 10.1016/j.ailsci.2021.100010.

Misra, N.N., Dixit, Y., Al-Mallahi, A., Bhullar, M.S., Upadhyay, R. and Martynenko, A. (2022), "IoT, Big Data, and Artificial Intelligence in Agriculture and Food Industry", IEEE Internet of Things Journal, Vol. 9 No. 9, pp. 6305–6324, doi: 10.1109/jiot.2020.2998584.

Mysona, D. P., Kapp, D. S., Rohatgi, A., Lee, D., Mann, A. K., Tran, P., ... & Chan, J. K. (2021). Applying artificial intelligence to gynecologic oncology: a review. Obstetrical & Gynecological Survey, 76(5), 292-301.

Nishant, R., Kennedy, M. and Corbett, J. (2020), "Artificial intelligence for sustainability: Challenges, opportunities, and a research agenda", International Journal of Information Management, Vol. 53, p. 102104, doi: 10.1016/j.ijinfomgt.2020.102104.

Norozpour, S. (2021). Simulation of the Relation Between the Number of COVID-19 Death Cases as a Result of the Number of Handwashing Facilities by Using Artificial Intelligence. Artificial Intelligence for COVID-19, 1-10.

Paton, C., & Kobayashi, S. (2019). An open science approach to artificial intelligence in healthcare. Yearbook of medical informatics, 28(01), 047-051.

Qazi, S., Khawaja, B.A. and Farooq, Q.U. (2022), "IoT-Equipped and AI-Enabled Next Generation Smart Agriculture: A Critical Review, Current Challenges and Future Trends", IEEE Access, Vol. 10, pp. 21219–21235, doi: 10.1109/access.2022.3152544.

Raclin, T., Price, A., Stave, C., Lee, E., Reddy, B., Kim, J., & Chu, L. (2022). Combining Machine Learning, Patient-Reported Outcomes, and Value-Based Health Care: Protocol for Scoping Reviews. JMIR Research Protocols, 11(7), e36395.

Rathinam, A. K., Lee, Y., Ling, D. N. C., Singh, R., Selvaratnam, L., & Pamidi, N. (2021, July). Artificial Intelligence in Medicine: A review of challenges in implementation and disparity. In 2021 IEEE International Conference on Health, Instrumentation & Measurement, and Natural Sciences (InHeNce) (pp. 1-6). IEEE.

Richards, B.A., Lillicrap, T.P., Beaudoin, P., Bengio, Y., Bogacz, R., Christensen, A., Clopath, C., et al. (2019), "A deep learning framework for neuroscience", Nature Neuroscience, Vol. 22 No. 11, pp. 1761–1770, doi: 10.1038/s41593-019-0520-2.

Sapienza, M., Nurchis, M. C., Riccardi, M. T., Bouland, C., Jevtić, M., & Damiani, G. (2022). The Adoption of Digital Technologies and Artificial Intelligence in Urban Health: A Scoping Review. Sustainability, 14(12), 7480.

Sarker, I.H. (2022), “AI-Based Modeling: Techniques, Applications and Research Issues Towards Automation, Intelligent and Smart Systems”, SN Computer Science, Vol. 3 No. 2, p. 158, doi: 10.1007/s42979-022-01043-x.

Straw, I., & Callison-Burch, C. (2020). Artificial Intelligence in mental health and the biases of language based models. PloS one, 15(12), e0240376.

Värri, A. (2020). What is digital health? Review of definitions. Integrated Citizen Centered Digital Health and Social Care: Citizens as Data Producers and Service co-Creators, 275, 67.

Wu, H., Lu, X., & Wang, H. (2023). The Application of Artificial Intelligence in Health Care Resource Allocation Before and During the COVID-19 Pandemic: Scoping Review. JMIR AI, 2(1), e38397.

Wang, W., Ye, Z., Gao, H. and Ouyang, D. (2021), “Computational pharmaceutics - A new paradigm of drug delivery”, Journal of Controlled Release, Vol. 338, pp. 119–136, doi: 10.1016/j.jconrel.2021.08.030.

Wille, S.M.R. and Elliott, S. (2020), “The Future of Analytical and Interpretative Toxicology: Where are We Going and How Do We Get There?”, Journal of Analytical Toxicology, Vol. 45 No. 7, pp. 619–632, doi: 10.1093/jat/bkaa13.

8 Appendix

8.1 Detailed analysis of REF Impact Case Studies

(Source: *REF Impact Case Studies*)

This taxonomy details the field of STEM, a summary of the data sources used in three UK REF ICS examples, a description of the AI related techniques described, an outline of the application domains/ commercial focus and additional examples with a list of other relevant impact case studies.

We examined how AI technology features in research practice, dissemination and impact by using the UK REF ICS database. We first identified 202 Impact Case Studies from the https://impact.ref.ac.uk/casestudies/ across units of assessment using keyword search “artificial intelligence” for summary of impact, underpinning research, references to the research and details of the impact. Then we filtered across main panels A, B and C. Those outside of the remit, e.g. social sciences within C, were excluded. Filtering by 12 UOA.

The application area or domain is used to classify the cases, rather than the Unit of Assessment. Impact case studies use descriptor ICS,1,2 and 3.

To produce this summary of the key themes identified in the Impact Cases, the following subject categorisations were used:1

- Physical sciences: Chemical Engineering, Chemistry, Computer Science, Earth and Planetary Sciences, Energy, Engineering, Environmental Science, Material Science, Mathematics, Physics and Astronomy.
- Health sciences: Medicine, Nursing, Veterinary, Dentistry, Health Professions.
- Life sciences: Agricultural and Biological Sciences, Biochemistry, Genetics and Molecular Biology, Immunology and Microbiology, Neuroscience, Pharmacology, Toxicology and Pharmaceutics

8.1.1 Physical sciences summary

Use cases for top AI techniques and applications

Themes (source: Literature review)	**Example use** cases (source: ICS)	**Top subfields** (source: Scopus)
Massive growth in data from advances in sensors Use of AI to model physical phenomena Importance of data and knowledge infrastructures (data standards and sharing, training in AI techniques for domain specialists) Constructive interactions between 'native' AI fields and application	Energy and Environmental Modelling at Building and Urban Scale Intelligent Energy Management FINEX: A probabilistic expert system for forensic identification The Natural Language Toolkit "Impact of Machine-Learning based Visual Analytics”	1. Electrical and electronic engineering 2. Control and systems engineering 3. General engineering 4. Theoretical computer science 5. General physics and astronomy

domains, leading to advances in AI technologies There are data gaps relating to publicly accessible datasets for model training, benchmark datasets for model evaluation, and benchmarks for the collection and curation of new datasets. There are also training needs.	Impact of Machine-Learning based Visual Analytics	

8.1.2 Life sciences summary

Use cases for top AI techniques and applications.

Themes (source: Literature review)	**Example use** cases (source: ICS)	**Top subfields** (source: Scopus)
Advancements in the Internet of Things (IoT) have provided a range of sophisticated sensors and physical devices that can be analysed using supervised and unsupervised AI models. 'Smart machines' are informing human interventions. Advancements in computing capability and algorithms have allow the integration of big data, artificial intelligence, and multi-scale modelling techniques Ethics and interpretable AI, data access and goal alignment The sector particularly relies on the IoT, big data, neural networks, modelling and image recognition.	The Utopia Suite: realising semantic knowledge discovery and data linkage in the publishing and pharmaceutical industries Increasing society's capacity to tackle complex, socio-technical dilemmas User-trainable visual anomaly detection for quality inspection tasks in the food industry Connecting Emotionally With and Through Computers Automated object recognition and focussing for Medical Applications	1. Cognitive neuroscience 2. General neuroscience 3. Biochemistry 4. General biochemistry, genetics and microbiology 5. Biotechnology

8.1.3 Health sciences summary

Use cases for top AI techniques and applications.

Themes (source: Literature review)	**Example use** cases (source: ICS)	**Top subfields** (source: Scopus)
AI in medicine is intertwined with ethics and considerations for human judgement Use of AI to support clinical decision-making and diagnosis and detection Importance of AI techniques to protect patient rights and participation	Biometrics: Supporting technology, policy and professional developments Using Product Design Techniques to Improve the Lives of Reconstructive Surgery Patients while Reducing the Cost to the NHS and Tax Payer	1. Health informatics 2. Radiology nuclear medicine and imaging 3. General medicine 4. Oncology 5. Medicine (misc.)

AI is transforming disease prediction, detection, modelling and diagnosis Use of robotics in healthcare presents HCI challenges	Automated object recognition and focussing for Medical Applications Robotics Applications in Health, Education and Entertainment Improved Mobility and Quality of Life for Children with Disabilities	

8.1.4 Detailed case study analysis

Below are relevant case studies, with their impact description extract from the UK REF Impact Case Study database. The tables outline the Case Study highlighted and a summary of the impact derived from the REF database.

8.1.4.1 Life science impact case studies

Case study	Impact
Intelligent Energy Management Unit of Assessment: Computer Science and Informatics	The work has had an economic, environmental and societal impact: it has shaped R&D strategies of leading British companies like BAE Systems and Secure Meters; the launch of iPhone apps and websites have supplied private and industrial users with personalized data regarding their energy use, resulting in cost savings and reductions in carbon emissions; it has enabled charities to provide energy-saving advice to households directly; and has won an international technology showcase competition leading to a spinout and commercialisation of research.
Energy and Environmental Modelling at Building and Urban Scale Unit of Assessment: Architecture, Built Environment and Planning	Developed advanced computational numerical models for simulating the energy and environmental performance of the built environment. These models have been used by leading design practices in the design of major buildings and urban developments. This impact case study presents three models from this research activity that have been widely taken up by industry worldwide, namely, the `building energy' model HTB2, the urban scale `energy and environment prediction' framework EEP and the `building environment' model ECOTECT.
Computer Based Methods for Diagnosing and Predicting River Health Unit of Assessment: Computer Science and Informatics	Computer-based solutions for the assessment of river water quality by environmental agencies, working to improve the quality. This research has informed discussions and decisions of the UK Technical Advisory Group for the Water Framework Directive (UKTAG WFD). UKTAG WFD have selected the WHPT (Walley, Hawkes, Paisley & Trigg) method, for assessing river water quality throughout the UK, in the context of river management to meet the targets set in the Water Framework Directive (Directive 2000/60/EC from the European Union), which the UK government signed up to in 2000.
New Tidal Flood Forecasting Systems Unit of Assessment: General Engineering	A new coastal flooding forecasting system combines forecasts of weather and sea conditions with modelling of wave transformation close to the coast, and from this information, using the outcomes of research at University of Liverpool between 1998 and 2005, predicts the wave overtopping of seawalls. The new system allows wind and wave conditions to be incorporated into coastal flooding predictions, improving on the previous methodology that was largely based on sea level. The Liverpool contribution to the system specifically improves on the conservatism of the previous overtopping prediction, leading to a model which issues less false alerts. Versions of the system are now in operation on the North East coast of England, and around the Firths or Forth and Tay, and over 200 alerts have been issued from the North East system since 2008.
The impact of lighting research into provision of user control and alternative daylight sources	The Lighting Group has been involved in the formulation of national and international design guidance, with impacts on the practices of the UK and international lighting industry.

Unit of Assessment: Architecture, Built Environment and Planning	This guidance offers designers the tools to create optimum visual conditions in energy efficient buildings while reducing electric lighting usage. This involves three areas: the development of lighting design in interiors to take account of room contents; involvement of the occupant in the control of light levels and electricity usage; and the replacement of electric lighting with guided daylight systems. The main impact of the work is its influence on the body of professional practice relating to interior lighting design. This guidance advocates the creation of user-friendly visual conditions, low electricity usage and natural light in areas remote from windows.
A sea-change in geophysical-marine surveying for protecting our Oceans future Unit of Assessment: Geography, Environmental Studies and Archaeology	The innovations in technology pioneered by this work also are providing critical findings on climate change impacts in the Earth's most sensitive and threatened environments with world media coverage on work in the Arctic including the award- winning TV series Operation Iceberg in 2012. Strong international media involvement has become one of the hallmarks of this work which simultaneously delivers research results as outputs of high quality across the globe. Furthermore, the technology has had economic impact in the form of three spin-out companies.
User-trainable visual anomaly detection for quality inspection tasks in the food industry Unit of Assessment: Computer Science and Informatics	A new multi-purpose computer vision system to identify sub-standard food products has been created. The research developed a user-trainable software technology with a range of possible applications, thus overcoming the specificity and other limitations such as the high set-up cost of existing visual inspection systems. This research is achieving impact in several areas within the food industry, including quality analysis of fresh produce, food processing and food packaging. The technology is currently being trialled at the leading post-harvest applied research facility for agricultural storage in the UK, and is also being licensed to a world-leading supplier of food packaging machines and equipment for inclusion in a new product range under development. The longer-term impacts include safer food, reduced food waste, more efficient food production, and better use of natural resources (e.g. reduced use of water, pesticides and other inputs), through early detection of potentially harmful flaws in production and packaging.

8.1.4.2 Health science impact case studies

Case study	**Impact**
New sensors for detecting oxygen levels in organs and tissues in critically ill patients Unit of Assessment: General Engineering	Research undertaken at City University London has led to the development of new blood oxygen optical and fibre optic sensors that advance clinical assessment in hospitals by monitoring a patient's arterial blood oxygen in specific organs or tissues. The applications of such sensors extend the boundaries of current state-of-the-art medical sensors in this field. They are capable of monitoring blood perfusion at times where the current commercial techniques fail to do so and have the advantage of providing organ-specific perfusion (oesophagus, bowel, liver, stomach, brain, etc.), enabling the effective monitoring of the wellbeing of specific parts of the body. These new sensors help clinicians monitor more reliably and provide the most appropriate treatment for very sick patients.
Using Product Design Techniques to Improve the Lives of Reconstructive Surgery Patients while Reducing the Cost to the NHS and Tax Payer Unit of Assessment: Art and Design: History, Practice and Theory	Over the last 15 years the Medical Applications Group (MAG) has engaged in applied research into the use of product design techniques and technologies in medical procedures. Their work has directly led to better, safer, faster, more accurate and less intrusive surgical procedures. The group has worked with surgeons at NHS hospitals all over the UK to deliver well over 2,000 medical models for surgical use during the period. A number of hospitals have adopted MAG's techniques, meaning that the Group's research has improved the dignity, comfort and quality of life of around two and a half thousand people since 2008 whilst saving the UK tax payer many thousands of pounds.
A new range of outdoor clothing for the active ageing based on wearable technologies	Two leading manufacturers of clothing for outdoor activities ([text removed for publication]) have produced a new range of functional clothing based on research at Ulster on wearable technologies for the active ageing. The new age-appropriate outdoor garments incorporate wearable technologies that enable self-monitoring of physiological parameters

Unit of Assessment: Computer Science and Informatics	(heart rate, respiration rate) and activity levels (step-counts, distance walked) with optimal placement of sensors to improve signal-to-noise ratio. Additionally, [text removed for publication], a company producing [text removed for publication], have used feedback from Ulster's research evaluations to design a new range of [text removed for publication] that are incorporated into the garments, achieving increased levels of usability by elderly people.
Walk This Way: Leading the World in Gait Biometrics Unit of Assessment: Electrical and Electronic Engineering, Metallurgy and Materials	Gait recognition research has produced impacts on public policy, on national security processes, on forensic service practice, on culture and society. The notion that people can be recognised by the way they walk was invented as a totally new means to identify people and has gained increasing popularity, reflected by its inclusion in an episode of BBC premier series Spooks. This followed considerable scientific development after its invention at Southampton in 1994, culminating in impacts that include its integration in a commercial system piloted by the National Physics Laboratory, novel forensic use in a criminal conviction, its take up by researchers at the Serious Organised Crime Agency and its focus by The Forensic Science Society. Southampton has retained its position at the forefront of gait biometrics research, collaborating nationally and internationally and driving prolific media engagement that has furthered this new technology and increased its global impact
Robotics Applications in Health, Education and Entertainment Unit of Assessment: Computer Science and Informatics	The Centre for Robotics and Neural Systems (CRNS) uses its research to address societal challenges, both nationally and internationally. It notably responds to practical problems and evaluates its robotics research in the real world, exposing it to use and users beyond the lab. This has generated both economic and social impact in clinical practice, education, entertainment and outreach: the use of robot companions for patients and disabled users; inspiration of school-children; engagement of thousands with the possibilities of robotics through high-profile robot competitions. Economic impact is reflected by commercial investment, and world-wide sales of robotics technologies by spin-off companies.
Interfacing brains with machines: public engagement and potential to benefit human health and quality of life Unit of Assessment: Electrical and Electronic Engineering, Metallurgy and Materials	The Cybernetics team at the University of Reading works at the frontier of human-machine interaction. The group carries out research on therapy and human enhancement in collaboration with medical professionals, to research new therapeutic treatments for patients with paralysis. Our work has led to the first human implantation of BrainGate, an intelligent deep brain stimulator, and the culturing of neurons within a robot body. Our work has been used by neurosurgeons in experimental human trials with the aim to enhance the standard of living of paralysed individuals. This ground-breaking, and sometimes controversial work, has sparked widespread discussion and debate in the public sphere, within the media and at the government level, on the use of machines to enhance humans and vice versa.
Connecting Emotionally With and Through Computers Unit of Assessment: Psychology, Psychiatry and Neuroscience	Emotional signals — obvious outbursts, or more often subtle changes in tone of voice, or facial expression — play a key part in human communication. Psychology researchers at Queen's have made fundamental contributions to 'affective computing', which enables automatic systems to use those signals. The team's work has influenced a new computing language for describing these signals and the states that they reveal: EmotionML (Emotion Markup Language). The language has been recommended as a standard by the World Wide Web Consortium, to define how software describes emotions. The language is used by multinational corporations in a range of applications in a rapidly expanding field. Queen's expertise in emotion led Dr Gary McKeown to found a start-up company, Adoreboard (previously known as Mediasights) along with entrepreneur Chris Johnston, which specifically uses EmotionML in opinion and sentiment analysis in marketing. Its product, Adoreboard, lets companies track consumers' emotional responses to their products. The company has agreed funding of £470,000, partnerships with three multinational corporations, and was recently selected to take up residence at Google's campus in London.
Automated object recognition and focussing for Medical Applications	This Keele University research in multiscale object recognition has led to two key breakthroughs: (a) the automated identification of tissue boundaries in computer tomographic (CT) scans, enabling the latest radiotherapy equipment to more accurately target diseased tissue thus avoiding neighbouring healthy organs. Such improvements are

Unit of Assessment: Computer Science and Informatics	essential to the successful roll-out of new more precise linear accelerators in the treatment of cancer; (b) new fractal algorithms to characterise the quality of transplanted cell growth from post-operative biopsies. By automating the selection of the healthiest cells this has assisted the generation of patient-specific cartilage and is essential for the development of a medical capability for large-scale patient-specific generation of cartilage growth for the treatment of arthritis. It has indirectly led to software improvements in cell-tracking and to achieving reliable auto-focussing in high throughput non-invasive microscopy.
Improved Mobility and Quality of Life for Children with Disabilities Unit of Assessment: Computer Science and Informatics	Research at the University of Portsmouth (UoP) has created new user-friendly control, navigation and communication systems for powered wheelchairs that have made a significant and positive impact on the lives of users. These have given many disabled children and adults an opportunity for independent mobility, some for the first time. The systems have been used in six special schools and institutions (including RNIB and NHS) and many private homes. Economic impact in reducing the need for carers alone has been estimated at more than £250,000 p.a and the devices have also changed some professional services.
Automated object recognition and focussing for Medical Applications Unit of Assessment: Computer Science and Informatics	This Keele University research in multiscale object recognition has led to two key breakthroughs: (a) the automated identification of tissue boundaries in computer tomographic (CT) scans, enabling the latest radiotherapy equipment to more accurately target diseased tissue thus avoiding neighbouring healthy organs. Such improvements are essential to the successful roll-out of new more precise linear accelerators in the treatment of cancer; (b) new fractal algorithms to characterise the quality of transplanted cell growth from post-operative biopsies. By automating the selection of the healthiest cells this has assisted the generation of patient-specific cartilage and is essential for the development of a medical capability for large-scale patient-specific generation of cartilage growth for the treatment of arthritis. It has indirectly led to software improvements in cell-tracking and to achieving reliable auto-focussing in high throughput non-invasive microscopy.
Increasing society's capacity to tackle complex, socio-technical dilemmas Unit of Assessment: Computer Science and Informatics	Compendium software is used to map dialogue and information around socio-technical dilemmas with economic, public policy, educational and health impacts. In Australia, urban planners attribute stakeholder buy-in to dialogue mapping with Compendium. In the USA, a deadlocked environmental planning process used it to make progress, while Southern California Edison use it to manage environmental policy. In the NHS, it can map therapeutic group dynamics, while in Germany, a journalist summarised a medical ethics case to parliament with it. More than 170 companies and individuals have endorsed Compendium, a striking application being to control Attention Deficit Hyperactivity Disorder (ADHD) at work.
The Utopia Suite: realising semantic knowledge discovery and data linkage in the publishing and pharmaceutical industries Unit of Assessment: Biological Sciences	The need to manage, analyse and interpret the volumes of data and literature generated by modern high-throughput biology has become a major barrier to progress. Research at the University of Manchester on interoperability and advanced interfaces has resulted in innovative software (Utopia Documents) that links biomedical data with scientific literature. The software has been adopted by international publishing houses (Portland Press, Elsevier, Springer, etc.), allowing them to explore new business models, and by pharmaceutical companies (e.g. AstraZeneca, Roche), providing new opportunities to explore more efficient, cost-effective methods for exploiting and sharing in-house data and knowledge. The research also led to a spin-out company, Lost Island Labs, in 2012, which expects a profit [text removed for publication] in its first year.

8.1.4.3 Physical science impact case studies

Case study	**Impact**
Engineering Knowledge for Autonomous and Intelligent Systems	University of Huddersfield research into knowledge engineering, domain modelling and machine learning has raised professional, industry and policymaker awareness of novel ways of designing more efficient, cost-effective and sustainable management networks. This is particularly the case in the field of transportation, where recognition of such

Unit of Assessment: Computer Science and Informatics	techniques has significantly increased among stakeholders throughout the UK and across Europe. The research has been credited with informing a "step-change in thinking" and is now central to the £16m EPSRC Autonomous and Intelligent Systems Programme, which has attracted more than £4m in financial and in-kind support from hi-tech industries
Data2Text Unit of Assessment Computer Science and Informatics	Data-to-text utilises Natural Language Generation (NLG) technology that allows computer systems to generate narrative summaries of complex data sets. These can be used by experts, professionals and managers to better, and quickly, understand the information contained within large and complex data sets. The technology has been developed since 2000 by Prof Reiter and Dr Sripada at the University of Aberdeen, supported by several EPSRC grants. The Impact from the research has two dimensions. As economic impact, a spinout company, Data2Text (www.data2text.com), was created in late 2009 to commercialise the research. As of May 2013, Data2Text had 14 employees. Much of Data2Text's work is collaborative with another UK company, Arria NLG (www.arria.com), which as of May 2013 had about 25 employees, most of whom were involved in collaborative projects with Data2Text. As impact on practitioners and professional services, case studies have been developed in the oil & gas sector, in weather forecasting, and in healthcare, where NLG provides tools to rapidly develop narrative reports to facilitate planning and decision making, introducing benefits in terms of improved access to information and resultant cost and/or time savings. In addition, the research led to the creation of simplenlg (http://simplenlg.googlecode.com/), an open-source software package which performs some basic natural language generation tasks. The simplenlg package is used by several companies, including Agfa, Nuance and Siemens as well as Data2Text and Arria NLG.
Applications of Singularity Theory and 3D Modelling in Arts and Retail Unit of Assessment: Mathematical Sciences	Professor Peter Giblin (Department of Mathematical Sciences at the University of Liverpool), together with collaborators, used methods from singularity theory to develop an approach for recovering 3-d information from 2-d images, such as photos. In the past decade, these have been implemented and built upon by software engineers, leading to significant cultural, economic and societal impacts. These include the creation of an innovative 25m high sculpture of the human body in the Netherlands by the sculptor Antony Gormley and the virtual modelling of clothing on online clothing websites such as Tesco's (Virtual Changing Room by Tesco/F&F). These have reached thousands of consumers worldwide and represent a significant commercial success for the company which developed the software.
Feature Recognition for Smart Design and Manufacture Unit of Assessment: General Engineering	In 1997 ERPE invented a novel automatic machining feature recognition technology which has been incorporated into the Pathtrace EdgeCAM Solid Machinist Computer Aided Manufacture (CAM) package, now owned by Planit plc. EdgeCAM is considered as one of the leading independent solid machinist CAM package, with 10 - 15% of the world market. Related ERPE feature recognition in shape representation and characterisation has enabled the design of a 3D shape browser for product data management systems. Commercialised in 2005 as Shapes pace with £0.7M current market value, for application to the parts industry in automotive markets, it has attracted the US Actify Inc., as an equity sharing partner to aid ShapeSpace to access worldwide markets.
Applications of agent technology Unit of Assessment: Computer Science and Informatics	Agent-based computing is a new paradigm for building complex socio-technical systems composed of many interacting intelligent and autonomous components. New coordination and negotiation algorithms developed at the University of Southampton, have provided new methods for managing such interactions in a flexible manner. This study focuses on their applications in two new start-up companies (Aerogility and Aroxo) in the defence, aerospace and civil contingency sectors (e.g. BAE Systems, Ministry of Defence and Hampshire County Council) in helping the GB Sailing Team to success at the 2012 Olympics, and in monitoring the environment for effects of climate change.
GRANIT Unit of Assessment: General Engineering	The GRANIT system is a non-destructive technique for assessing the condition of rock bolts and ground anchors used to support structures such as tunnels. It applies a small impulse to the bolt and interprets the resulting vibration response to provide estimates of load and unbonded length. Initial development of the system was based on the findings of

	EPSRC projects in tunnels undertaken by the Universities of Aberdeen and Bradford from 1989-1997, resulting in an empirically based method. However, research undertaken at the University of Aberdeen since 1998 has provided the understanding of the process and developed the fundamental engineering science needed to underpin the development of a full commercial system. The GRANIT system is patented, and has been subject to worldwide licence to Halcrow who have undertaken testing and provided a method of ensuring the safety of mines, tunnels and similar structures. Halcrow received the NCE award for Technical Innovation Award for GRANIT in December 2010. The impact of the research has been in part economic, but largely on practitioners and professional services.
Classification within forensic datasets Unit of Assessment: Computer Science and Informatics	This Keele University research into advanced signal processing and classification methods has led to novel algorithms capable of isolating subtle patterns in complex data. This has been applied in two highly significant application areas: first to the problem of image source identification and second to the problem of unobtrusive but highly secure authentication methods. In the first case this has enabled images captured by mobile phone cameras to be reliably and evidentially linked to source devices. This has huge applicability to those fighting terrorism, paedophile rings and civil unrest by extending detection capabilities to mobile phones in an era in which they are rapidly replacing dedicated cameras. It helps to prove, for example, that a photograph entered as evidence was captured by a specific mobile phone. As most phones can be tied to their user or owner this is extremely important to the successful detection and prosecution of offenders. In the second case it has enabled criminal record checks to be carried out securely online where previous paper-based systems were both too slow for purpose (taking weeks or months) and inherently insecure, leaving key posts unfilled in the healthcare industries and education sector; so, benefiting the public by solving a problem that was having a negative impact on the running of these public services.
Development of Smart Planning Tools for BT and Network Optimisation Unit of Assessment: Computer Science and Informatics	A series of funded research projects have been completed by the University of Sunderland in close collaboration with BT Research Labs Ipswich. This research, which has resulted in a series of novel optimisation approaches, led to the development of a suite of tools used for network planning. These tools are primarily based upon the application of evolutionary computing methods. Researchers produced intelligent network planning tools for the development of the national Internet. The tools have been used extensively since 2008, and the network for the Olympic games in London 2012 was designed and planned to use these smart tools. A company specialising in vehicle tracking has also been formed as a direct result of the research.
Controlling uncertainty with cost engineering tools Unit of Assessment: Aeronautical, Mechanical, Chemical and Manufacturing Engineering	Substantial savings have been made using Cranfield's Cost Engineering software tools and techniques. These are used in BAE Systems, Airbus, Rolls-Royce, GE Aviation, Ford Motor Company and increasingly in the UK defence industry through the MoD. DTZ (Debenham Tie Leung Ltd) estimates £213 million per annum financial benefit for BAE Systems and MoD alone, with an additional £200 million per annum for other companies. Cranfield's team has significantly influenced the national and international policy of The Association of Cost Engineers and manufacturing companies in methods and procedures. Cranfield has trained over 700 engineers from over 50 companies in cost engineering based on our research.
EFIT-V Facial Recognition Software Unit of Assessment: Physics	Research conducted within the School of Physical Sciences (SPS) at the University of Kent has led to the development and successful commercialisation of facial identification software named EFIT-V. First sold in 2007, this software is now used by more than 70 police forces internationally and has revolutionised the way eyewitnesses and victims of crime create computerised facial likenesses of offenders. These images are circulated to police intelligence units, and the general public, leading to the identification and arrests of offenders. Police Identification rates have jumped from 5% to 55% as a result of this software. With a current annual turnover exceeding £250K, which is projected to reach £600K by 2015, Kent spinout company Visionmetric has made significant impact with EFIT-V, and achieved a position of commercial dominance in the UK, and around the world.

The Natural Language Toolkit (NLTK) Unit of Assessment: Computer Science and Informatics	The Natural Language Toolkit (NLTK) is a widely-adopted Python library for natural language processing. NLTK is run as an open-source project. Three project leaders, Steven Bird (Melbourne University), Edward Loper (BBN, Boston) and Ewan Klein (University of Edinburgh) provide the strategic direction of the NLTK project. NLTK has been widely used in academia, commercial / non-profit organisations and public bodies, including Stanford University and the Educational Testing Service (ETS), which administers widely-recognised tests across more than 180 countries. NLTK has played an important role in making core natural language processing techniques easy to grasp, easy to integrate with other software tools, and easy to deploy.
An Innovative Friction Welding Platform for Creep Damage Assessment and Repair of Thermal Power Plant Components Unit of Assessment: General Engineering	This case study deals with research undertaken at Plymouth University leading to the development of an innovative friction stir welding process (friction hydro-taper pillar processing, FHPP) and a bespoke welding platform that improves the assessment and repair methodology for creep damaged thermal power station components. This technology, developed in collaboration with Nelson Mandela Metropolitan University and with industry investment, enables power station engineers to extend the life of power generating plant leading to multi-million pound cost savings (over £66M in direct financial savings are demonstrated in this case) plus significant safety and societal impacts. It has been patented in South Africa and a spin-off company has been formed.
Making sense of complex data through innovations in visualisation Unit of Assessment: Computer Science and Informatics	New visualisation approaches have been used to turn complex data into actionable knowledge by: - The Ministry of Defence to establish analytical possibilities for security critical data analysis - Transport for London (TfL) to manage and extend London's successful Cycle Hire Scheme - E.ON to interpret data produced through their modelling and in their Smart Home trial, with a view to understanding electricity consumption and reducing production - Leicestershire County Council (LCC) to develop an evidence base for a sustainable transport plan; record and analyse the locations of locally valued green spaces; capture local knowledge about flooding events; monitor performance of children's centres; present the results of a survey on service quality and accessibility to citizens; undertake a £100M budget consultation and embed data in decision-making processes to inform policy - Willis to understand and assess windstorm risk, communicate the complexity of risk to clients and manage risk across their global offices through a new software system. These applications of new visualisation methods have had an impact on the environment, economy, defence and security, society and public debate. In each case users of our methods report on their positive impact as we help them identify visualisation possibilities, understand their data and use this knowledge to inform their activity. In many cases our work has resulted in important insights, improved exploitation of data and further investment in visualisation with organisational implications in terms of using data for intelligence.
Impact of Machine-Learning based Visual Analytics Unit of Assessment: Computer Science and Informatics	Visual analytics is a powerful method for understanding large and complex datasets that makes information accessible to non-statistically trained users. The Non-linearity and Complexity Research Group (NCRG) developed several fundamental algorithms and brought them to users by developing interactive software tools (e.g. Netlab pattern analysis toolbox in 2002 (more than 40,000 downloads), Data Visualisation and Modelling System (DVMS) in 2012). Industrial products. These software tools are used by industrial partners (Pfizer, Dstl) in their business activities. The algorithms have been integrated into a commercial tool (p:IGI) used in geochemical analysis for oil and gas exploration with a 60% share of the worldwide market.

	Improving business performance. As an enabling technology, visual analytics has played an important role in the data analysis that has led to the development of new products, such as the Body Volume Index, and the enhancement of existing products (Wheelright: automated vehicle tyre pressure measurement). Impact on practitioners. The software is used to educate and train skilled people internationally in more than 6 different institutions and is also used by finance professionals.
Sentic Computing Unit of Assessment: Computer Science and Informatics	Extracting information and meaning from natural language text is central to a wide variety of computer applications, ranging from social media opinion mining to the processing of patient health-care records. Sentic Computing, pioneered at the University of Stirling, underpins a unique set of related tools for incorporating emotion and sentiment analysis in natural language processing. These tools are being employed in commercial products, with performance improvements of up to 20% being reported in accuracy of textual analysis, matching or even exceeding human performance (Zoral Labs). Current applications include social media monitoring as part of a web content management system (Sitekit Solutions Ltd), personal photo management systems (HP Labs India) and patient opinion mining (Patient Opinion Ltd). Impact has also been achieved through direct collaboration with other commercial partners such as Microsoft Research Asia, TrustPilot and Abies Ltd. Moreover, international organisations such as the Brain Sciences Foundation and the A*Star Institute for High Performance Computing have realised a major impact by drawing upon our research.
FINEX: A probabilistic expert system for forensic identification Unit of Assessment: Mathematical Sciences	The mathematical calculations for determining the likelihood of kinship between two individuals from their DNA profiles can be quite laborious and error prone, even when carried out by experts in the field. FINEX, an Expert System-based software programme to automate such calculations, was developed through research in Bayesian networks undertaken at City University London. The software can accurately and efficiently calculate kinship likelihoods within a minute or two, calculations that could take an expert half a day or more. The software was licensed to the UK Forensic Science Service (FSS), who used it among other applications to analyse DNA evidence leading to convictions for several serious and high-profile criminal cases.
Metail Unit of Assessment: General Engineering	Research at the University of Cambridge Department of Engineering (DoEng) since 1997 created methods for reconstructing a three-dimensional (3D) model of an object from a single two-dimensional photograph. Metail, a company founded in 2008, sponsored further research at the DoEng and commercialised the results in an online fashion retailing application. Metail enables customers to select an item of clothing and see how they would look wearing it from a variety of angles, having entered just one photograph of themselves and a few basic body measurements. Metail attracted over GBP3.5M investment. Its application is used by Shop Direct, Tesco, Warehouse, Zalando and Dafiti. Sales data shows that the Metail application increases the propensity of customers to buy and reduces the proportion of goods returned.
Economic benefits from sales of people-tracking and crowd-monitoring technology Unit of Assessment: Computer Science and Informatics	Research at Kingston University into methods for tracking pedestrians and monitoring crowds using computer vision techniques has been translated into commercial products by Ipsotek Ltd and BAe Systems, resulting in economic benefits to these companies from sales of these products. These products have been sold to high-profile customers including the London Eye, the O2 Arena and the Australian Government, providing significant commercial benefits, employment and growth for both companies, as well as providing an economic impact for these customers.
VC Learning Theory and Support Vector Machines Unit of Assessment: Computer Science and Informatics	Results of research at Royal Holloway in Machine Learning — Support Vector Machine (SVM), kernel methods and conformal-prediction techniques — are at the source of the analytics and `Big Data' revolution, whose impact is transforming the economy (and society), from data mining to machine vision, from Web search to spam detection.

SpendInsight Unit of Assessment: Computer Science and Informatics	Bishop and Danicic contributed to the development of novel spend analysis software. Launched in 2011 as a commercial service by KTP industrial partners @UK PLC, SpendInsight has been used by over 380 organisations, including Basingstoke and North Hampshire NHS Foundation Trust, which, alone, cut procurement spend by £300,000 via savings identified using SpendInsight. An analysis produced by SpendInsight for the National Audit Office identified gross inefficiencies in NHS procurement, yielding potential annual overall savings of at least £500 million. The findings of this report were discussed in parliament and changes to NHS purchasing policy were recommended as a result.
Optical OFDM Transceiver Development and Commercialisation Unit of Assessment: Electrical and Electronic Engineering, Metallurgy and Materials	Pioneering research at Bangor on the advanced communications technology termed Optical Orthogonal Frequency Division Multiplexing (OOFDM) has enabled industrial impact with global implications. OOFDM was a candidate technique for the ITU-T G989.1 NG-PON2 and the IEEE 802.3bm standards and is currently under consideration by the IEEE 802.3 400Gb/s Ethernet Study Group. Supported by 8 patent families and first-phase funding of £1.1M, in 2013, the pre-revenue Bangor University spin-off company Smarterlight Limited, was established. Smarterlight has deployed services to several international telecommunications companies to develop advanced solutions for access optical networks and data centres.
Design of Authentication Algorithms for GSM Phones Unit of Assessment: Mathematical Sciences	Mobile telecommunication networks serve nearly 7 billion users; over 90% of the world's population. The flexibility and pervasive nature of mobile networks underpin an enormous range of business and personal activities. Many systems are based on GSM (Global System for Mobile Communications) standards for digital cellular networks that were created by the European Telecommunications Standards Institute (ETSI) in the 1990s to replace analogue network standards. A key factor in the success of GSM has been the ability to authenticate legitimate users and to provide privacy for wireless transmissions. A strong authentication mechanism is critical for the economic operation of mobile telephony. The security of GSM is based on a secret key, known only to the network operator and the Subscriber Identity Module (SIM), and an authentication algorithm implemented by the SIM and the network operator. A network operator may implement its own authentication algorithm, but many adopted the example implementation (known as COMP128, or COMP128-1) suggested by the GSM Association (GSMA). COMP128-1 was later found to be flawed. Cryptographers at Royal Holloway, at the request of GSMA, designed a replacement algorithm (COMP128-2), the example implementation offered by the GSM Association (GSMA) to over 800 Mobile Network Operators (MNO) in over 200 countries. The algorithm is still regarded as robust and it and derivative algorithms are relied upon by enormous numbers of users every day.
User Authentication Methodologies for Secure and Competitive Business Unit of Assessment: Computer Science and Informatics	Between 2003-2008 our research into an efficient multi factor-multimodal biometric authentication method for smartcards enabled Ecebs (http://www.ecebs.com/), a small-to-medium enterprise company specialising in smartcard software solutions to increase their patent portfolio, widen its product and service offering, improve their competitive position and create new business opportunities. In 2007 Ecebs was acquired by Trainline Investment Holdings Ltd (today known as The Trainline.com), and subsequently in 2012 by Bell ID, a global provider of smartcard/contactless/mobile solutions.
Biometrics: Supporting technology, policy and professional developments Unit of Assessment: General Engineering	Our impact on the theory and practice of biometrics (identification of individuals through measurement/analysis of their physiological/behavioural characteristics) embraces contributions to technological development, to general systems-level principles and to public policy and professionalisation issues. Our research and consequent engagement across the stakeholder community has impacted on the technological development of practical biometrics through take-up by industry (e.g. InMezzo, one of the UK's leading secure information specialists, has enhanced identity authentication procedures), company spinout (the EFIT-V facial recognition suite from VisionMetric Ltd fundamentally changed the means by which facial composites are created and is now used by more than 85% of Britain's Police Forces), leadership of the development of standards for the expanding commercial marketplace (e.g. establishment of standards for image acquisition for e-passports and other access control applications) and policy-level input to Government and International Professional Bodies, providing long-term support for practical deployment

	and end- user engagement (the Biometrics Assurance Group with Fairhurst as an independent member reported the security risk and problems identifying fingerprints within the UK government's £5.6bn ID card scheme proposal).
Intelligent Pricing Decision Support Systems Unit of Assessment: Computer Science and Informatics	Pricing optimisation and revenue management systems represent fundamental progress from the art of pricing to the science of pricing. Our research led the scientific approach in demand modelling and pricing optimization, and produced the first computerised Intelligent Pricing Decision Support Systems (IPDSS) for retail and petroleum, which have led to economic impact and changes in pricing practice. Our research led to spin-off companies that employ over 150 people, with a turnover of £19.2m in 2012, which are the leading providers of IPDSS, used by more than 400 retailers across 80 countries to improve their performance in competitive markets.
Kelvin Connect – a highly successful spin-out providing advanced mobile data capture systems for police officers and healthcare professionals Unit of Assessment: Computer Science and Informatics	A quiet technology revolution in the UK has been changing the way that police officers on the beat and hospital nurses access and record information, using handheld electronic notebooks that bring large time and cost savings. This revolution began as a University of Glasgow research programme and led to the creation of a successful spin-out company, Kelvin Connect. Acquired in 2011 by the UK's largest provider of communications for emergency services, Kelvin Connect has grown to 30 staff. Its Pronto systems are now in use by 10% of UK police forces and nursing staff in several UK hospitals.
Improved video surveillance and customer relations management through efficient data representation. Unit of Assessment: Computer Science and Informatics	Research on data compression produced novel algorithms that optimise the use of bandwidth and processing power. This research has led to the establishment of a product line that applies these algorithms to video surveillance software, marketed by Digital Barriers plc. Since 2008 this compression technology has allowed the company to grow from 8 to 41 staff and increase revenue from £800K to £6M in 2013. The novelty and usefulness of the data compression research was also appreciated by ThinkAnalytics plc. This led the company to the optimal design for data compression in their recommender system, which is currently being supplied to 130M cable TV customers making the product the most deployed content recommendation system in the market.
GATE: General Architecture for Text Engineering Unit of Assessment: Computer Science and Informatics	GATE (a General Architecture for Text Engineering—see http://gate.ac.uk/) is an experimental apparatus, R&D platform and software suite with a very wide impact in society and industry. There are many examples of applications: the UK National Archive uses it to provide sophisticated search mechanisms over its .gov.uk holdings; Oracle includes it in its semantics offering; Garlik Ltd. uses it to mine the web for data that might lead to identity theft; Innovantage uses it in intelligent recruiting products; Fizzback uses it for customer feedback analysis; the British Library uses it for environmental science literature indexing; the Stationery Office for value-added services on top of their legal databases. It has been adopted as a fundamental piece of web infrastructure by major organisations like the BBC, Euromoney and the Press Association, enabling them to integrate huge volumes of data with up-to-the-minute currency at an affordable cost, delivering cost savings and new products.

8.2 An attempt to identify underlying technologies used in AI research

The below table reports our early attempts to discern the underlying technologies used in AI research across different fields through the Scopus 'Topic name' schema. In this analysis, all AI publications from 2022 have been used. Each publication was allocated to a single ASJC code, based on the first code listed with the publication. From these codes, publications were allocated into fields and subfields. The frequency of Topic name categories within each subfield were then calculated. The table shows the 100 most frequently occurring Topic names. As can be seen, the ‘Object Detection,Deep Learning,IOU’ topic category is the most frequently occurring, across a wide range of subfields and fields.

Field	**Subfield**	**Topic name**	**Count**
Computer Science	Artificial Intelligence	Object Detection,Deep Learning,IOU	1933
Computer Science	General Computer Science	Object Detection,Deep Learning,IOU	1090
Computer Science	Computer Networks and Communications	Object Detection,Deep Learning,IOU	1024
Computer Science	Computer Science Applications	Object Detection,Deep Learning,IOU	903
Earth and Planetary Sciences	General Earth and Planetary Sciences	Object Detection,Deep Learning,IOU	327
Computer Science	Artificial Intelligence	Collaborative Filtering,Recommender Systems,Factorization	308
Computer Science	Computer Graphics and Computer-Aided Design	Object Detection,Deep Learning,IOU	269
Computer Science	Artificial Intelligence	Embedding,Named Entity Recognition,Entailment	260
Engineering	Electrical and Electronic Engineering	Object Detection,Deep Learning,IOU	252
Computer Science	Artificial Intelligence	Multiagent Learning,Multi-Agent Systems,Policy Iteration	244
Computer Science	Computer Networks and Communications	Embedding,Named Entity Recognition,Entailment	231
Computer Science	General Computer Science	Embedding,Named Entity Recognition,Entailment	222
Biochemistry, Genetics and Molecular Biology	Biochemistry	Object Detection,Deep Learning,IOU	199
Earth and Planetary Sciences	General Earth and Planetary Sciences	Hyperspectral Imagery,Spectroscopy,Image Classification	198
Agricultural and Biological Sciences	Agronomy and Crop Science	Object Detection,Deep Learning,IOU	177
Chemical Engineering	Fluid Flow and Transfer Processes	Object Detection,Deep Learning,IOU	169
Engineering	Industrial and Manufacturing Engineering	Object Detection,Deep Learning,IOU	154
Engineering	General Engineering	Object Detection,Deep Learning,IOU	127
Earth and Planetary Sciences	Geotechnical Engineering and Engineering Geology	Object Detection,Deep Learning,IOU	109
Engineering	Biomedical Engineering	Object Detection,Deep Learning,IOU	90
Earth and Planetary Sciences	General Earth and Planetary Sciences	Remote Sensing Image,Convolutional Neural Network,Image Retrieval	89
Engineering	Electrical and Electronic Engineering	Particle Accelerators,Convolutional Neural Network,TOPS	79
Materials Science	Biomaterials	Object Detection,Deep Learning,IOU	77
Energy	Energy Engineering and Power Technology	Battery Management Systems,Battery Pack,Charging (Batteries)	74

Engineering	Electrical and Electronic Engineering	Rolling Bearing,Rotating Machinery,Failure Analysis	73
Business, Management and Accounting	Management Information Systems	Object Detection,Deep Learning,IOU	71
Medicine	Radiology Nuclear Medicine and imaging	Object Detection,Deep Learning,IOU	69
Engineering	Mechanical Engineering	Rolling Bearing,Rotating Machinery,Failure Analysis	67
Earth and Planetary Sciences	Geotechnical Engineering and Engineering Geology	Hyperspectral Imagery,Spectroscopy,Image Classification	66
Physics and Astronomy	General Physics and Astronomy	Object Detection,Deep Learning,IOU	65
Engineering	Civil and Structural Engineering	Object Detection,Deep Learning,IOU	65
Medicine	Radiology Nuclear Medicine and imaging	Texture Analysis,Cancer,Fluorodeoxyglucose F 18	63
Environmental Science	Water Science and Technology	Prediction,Flood Forecasting,Water Tables	63
Engineering	Control and Systems Engineering	Object Detection,Deep Learning,IOU	63
Energy	Energy Engineering and Power Technology	Object Detection,Deep Learning,IOU	55
Biochemistry, Genetics and Molecular Biology	Cancer Research	Texture Analysis,Cancer,Fluorodeoxyglucose F 18	54
Agricultural and Biological Sciences	General Agricultural and Biological Sciences	Object Detection,Deep Learning,IOU	54
Energy	Energy Engineering and Power Technology	Wind Speed,Neural Networks,Prediction Interval	53
Earth and Planetary Sciences	Geotechnical Engineering and Engineering Geology	Remote Sensing Image,Convolutional Neural Network,Image Retrieval	53
Earth and Planetary Sciences	General Earth and Planetary Sciences	Landsat,Land Cover,Remote Sensing	52
Energy	General Energy	Wind Speed,Neural Networks,Prediction Interval	51
Engineering	Biomedical Engineering	Motor Imagery,Brain Computer Interface,Electroencephalogram	48
Earth and Planetary Sciences	General Earth and Planetary Sciences	Landslides,Debris Flow,Susceptibility	48
Earth and Planetary Sciences	Atmospheric Science	Object Detection,Deep Learning,IOU	48
Agricultural and Biological Sciences	Plant Science	Object Detection,Deep Learning,IOU	48
Earth and Planetary Sciences	Earth and Planetary Sciences (miscellaneous)	Object Detection,Deep Learning,IOU	46
Biochemistry, Genetics and Molecular Biology	General Biochemistry, Genetics and Molecular Biology	Object Detection,Deep Learning,IOU	46
Energy	Energy Engineering and Power Technology	Electric Power Plant Loads,Electricity Price,Power Markets	45
Biochemistry, Genetics and Molecular Biology	Biophysics	Object Detection,Deep Learning,IOU	45

Agricultural and Biological Sciences	Ecology, Evolution, Behavior and Systematics	Object Detection,Deep Learning,IOU	44
Agricultural and Biological Sciences	Aquatic Science	Prediction,Flood Forecasting,Water Tables	43
Energy	Energy Engineering and Power Technology	Diffuse Solar Radiation,Clear Sky,Prediction	42
Energy	General Energy	Electric Power Plant Loads,Electricity Price,Power Markets	40
Biochemistry, Genetics and Molecular Biology	Biochemistry	Rolling Bearing,Rotating Machinery,Failure Analysis	40
Chemistry	Physical and Theoretical Chemistry	Neural Networks,Potential Energy Surfaces,Materials Science	39
Arts and Humanities	Language and Linguistics	Embedding,Named Entity Recognition,Entailment	39
Biochemistry, Genetics and Molecular Biology	Clinical Biochemistry	Object Detection,Deep Learning,IOU	38
Energy	Renewable Energy, Sustainability and the Environment	Diffuse Solar Radiation,Clear Sky,Prediction	37
Medicine	Health Informatics	Electronic Medical Record,Natural Language Processing (NLP),Anonymization	36
Energy	General Energy	Diffuse Solar Radiation,Clear Sky,Prediction	36
Biochemistry, Genetics and Molecular Biology	Biotechnology	Object Detection,Deep Learning,IOU	36
Agricultural and Biological Sciences	Agronomy and Crop Science	Crops,Leaf Area Index,Hyperspectral Data	34
Chemical Engineering	General Chemical Engineering	Batch Process,Fault Detection,Canonical Variate Analysis	32
Agricultural and Biological Sciences	Food Science	Hyperspectral Imaging,Total Volatile Basic Nitrogen,Fruit	32
Environmental Science	Water Science and Technology	Landslides,Debris Flow,Susceptibility	31
Energy	General Energy	Battery Management Systems,Battery Pack,Charging (Batteries)	31
Energy	Energy (miscellaneous)	Electric Power Plant Loads,Electricity Price,Power Markets	31
Agricultural and Biological Sciences	Food Science	Object Detection,Deep Learning,IOU	31
Chemistry	General Chemistry	Neural Networks,Potential Energy Surfaces,Materials Science	30
Biochemistry, Genetics and Molecular Biology	Biochemistry	Activity Recognition,Wearable Sensors,Sensor	30
Medicine	Health Informatics	Radiological Findings,Clinical Features,COVID-19	29
Business, Management and Accounting	Management Information Systems	Embedding,Named Entity Recognition,Entailment	29
Agricultural and Biological Sciences	Aquatic Science	Object Detection,Deep Learning,IOU	28

Medicine	Radiology Nuclear Medicine and imaging	Radiography,Pneumothorax,Artificial Intelligence	26
Materials Science	General Materials Science	Object Detection,Deep Learning,IOU	26
Environmental Science	Environmental Chemistry	Prediction,Flood Forecasting,Water Tables	26
Business, Management and Accounting	Management of Technology and Innovation	Object Detection,Deep Learning,IOU	26
Biochemistry, Genetics and Molecular Biology	Biochemistry	Position Weight Matrix,Jackknife,Support Vector Machine	26
Medicine	Radiology Nuclear Medicine and imaging	Radiological Findings,Clinical Features,COVID-19	25
Chemical Engineering	General Chemical Engineering	Object Detection,Deep Learning,IOU	25
Chemical Engineering	Fluid Flow and Transfer Processes	Rolling Bearing,Rotating Machinery,Failure Analysis	25
Agricultural and Biological Sciences	Agronomy and Crop Science	Robot,End Effectors,Malus	25
Medicine	General Medicine	Radiological Findings,Clinical Features,COVID-19	24
Business, Management and Accounting	Management Information Systems	Collaborative Filtering,Recommender Systems,Factorization	24
Biochemistry, Genetics and Molecular Biology	General Biochemistry, Genetics and Molecular Biology	Position Weight Matrix,Jackknife,Support Vector Machine	24
Biochemistry, Genetics and Molecular Biology	Cancer Research	Object Detection,Deep Learning,IOU	24
Materials Science	General Materials Science	Neural Networks,Potential Energy Surfaces,Materials Science	21
Chemical Engineering	Fluid Flow and Transfer Processes	Navier-Stokes Equations,Model Form,Reynolds Stress	21
Arts and Humanities	Language and Linguistics	Speaker Adaptation,Acoustic Model,Deep Neural Network	21
Neuroscience	General Neuroscience	Connectome,Functional Magnetic Resonance Imaging,Functional Connectivity	20
Medicine	General Medicine	Object Detection,Deep Learning,IOU	19
Chemical Engineering	General Chemical Engineering	Neural Networks,Potential Energy Surfaces,Materials Science	19
Medicine	Health Informatics	Object Detection,Deep Learning,IOU	18
Chemistry	Physical and Theoretical Chemistry	Chemoinformatics,Drug Discovery,Topographic Mapping	18
Materials Science	Electronic, Optical and Magnetic Materials	Berry Phase,Holograms,Optics	17
Neuroscience	General Neuroscience	Spiking Neural Networks,Sensor,Event-Driven	16
Medicine	Pathology and Forensic Medicine	Breast Neoplasms,Cancer Classification,Histopathology	16
Chemical Engineering	Fluid Flow and Transfer Processes	Embedding,Named Entity Recognition,Entailment	16

Business, Management and Accounting	Strategy and Management	Object Detection,Deep Learning,IOU	16
Arts and Humanities	Language and Linguistics	Object Detection,Deep Learning,IOU	16