

Institution: University of Edinburgh

Unit of Assessment: 11 (Computer Science and Informatics)

1. Unit context and structure, research and impact strategy

1.1 Context, strategy and structure

Informatics at Edinburgh is based on a vision to support the **transformation of society through the generation, manipulation and understanding of information**, in our increasingly digital age. To bring about this vision we conduct foundational research within the discipline but also reach out across disciplinary and academic boundaries, working with others to ensure that the full potential of our research achievements is realised. This vision is central to the university's strategy, exemplified by the large investment in Data-Driven Innovation, which has attracted £237M to the University of Edinburgh, awarded by the UK and Scottish Governments within the £1.3B Edinburgh and South East Scotland City Region Deal (the "Edinburgh City Deal"), and resulting in the Bayes Centre, our new £45M innovation hub for Data Science and AI, opened in 2018, adjacent to our main home, the Informatics Forum (opened 2008).

To realise our vision we have grown considerably in the REF period with the number of category-A staff increasing by 52% to 144.14 FTE and our annual research awards increasing by 88% to £23.9M/year (REF-2021) from £12.7M/year (REF-2014). This has resulted in the development of new or substantially increased research activities, and new professorial appointments, in Biomedical AI, Cyber Security and Privacy, Digital Learning, and Quantum Computing. Our research excellence has been recognised by many prestigious **prizes and fellowships** during the REF period, including the BCS Needham Award, Lovelace Medal, and Spärck Jones lecture, and 2 new fellows of the Royal Society, the Royal Society of Edinburgh (3), the Royal Academy of Engineering (1), Academia Europaea (5), ACM (1), and the IEEE (1).

We have focused on seven strategic goals during the REF period:

- Attraction, development, and retention of **global talent** – 64 category-A staff appointed during the REF period (a net increase of 60 over REF-2014) and 68 promotions.
- A step change in **interdisciplinary and translational research** – focused on the Bayes Centre, which convenes cross-university groups to support new research and innovation activities, housing 650 staff and students, a third of whom are from industry partners.
- World-leading **research facilities** – notably the £82M ARCHER2 supercomputer (2020), and the £118M Edinburgh International Data Facility (2018), supported by the Edinburgh City Deal, and operated by EPCC, our supercomputing centre.

- Excellent **research training** – establishment of five UKRI CDTs, and an increase of 80% in our PGR student cohort, from 280 (268 PhDs and 12 MRes) on course in July 2013 to 504 (481 PhDs and 23 MRes) on course in July 2020.
- Diversification of **research funding** – research awards nearly doubled from £12.7M/year in REF-2014 to £23.9M/year now, with £9.9M/year from UKRI, £7.2M/year from H2020 and £22.6M from Industry (overview in Figure 3).
- Substantial **industrial research collaborations** – our industry research funding, representing a greater than 5 factor increase compared with the REF-2014 period, reflects wide engagement with industrial collaborators.
- Construction of **networks of collaboration** internationally, nationally, and locally – for example, we are founder members of the Alan Turing Institute (2015) and Health Data Research UK (2018); 20 of our category-A staff are Turing Fellows, and we have established five joint industry research labs with ARM, Huawei, IOHK, RBS, and Rolls-Royce.

Achieving these goals in the REF period has resulted in an environment that enables ground-breaking research with real-world impact across the discipline of computer science and informatics. Our research is structured across four themes – **Computer Systems, Data Science and Artificial Intelligence, Foundations of Computation, and Language, Interaction, and Robotics** – described below. We have also massively increased our interdisciplinary research activities. Informatics is at the core of the university's interdisciplinary activity, with active collaborations with 22 out of 24 of the schools in the university (Figure 1).

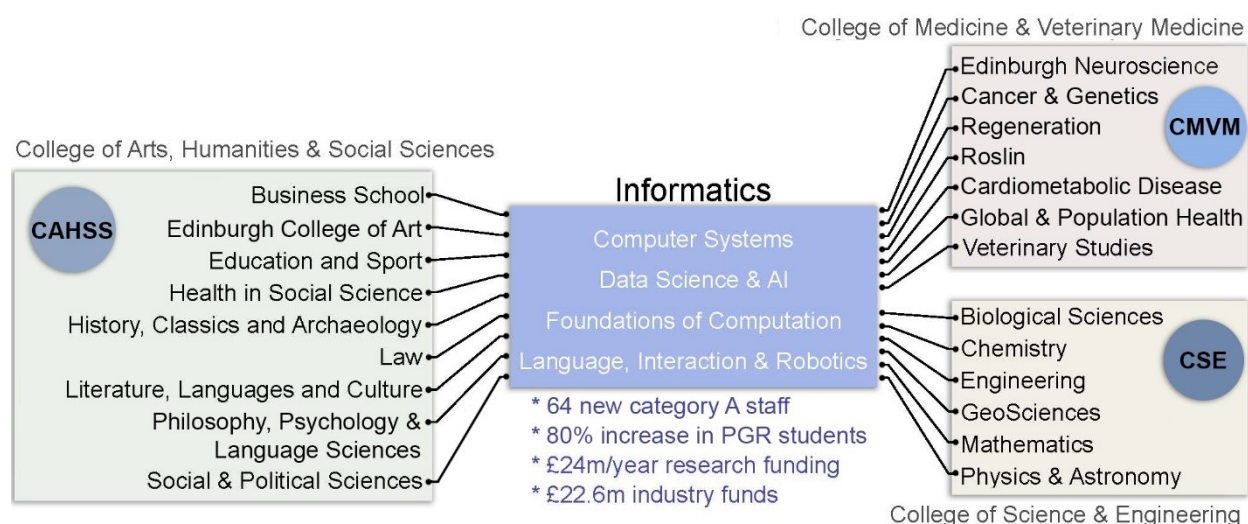


Figure 1: Interdisciplinary collaborations across the university.

1.2 Research Themes

The themes are presented in Subsections A – D; research awards by theme are presented in Section 3.1, and further impact arising from each theme in Section 4.2. In Subsection E we present our environment for interdisciplinary research.

A. Computer Systems

Theme leaders: **Nagarajan, Parsons**

28 category-A staff, £64.2M research awards; 57 current PhD students

16 new category-A staff in the REF period: **Ainsworth, Barbalace, Bhatotia, Brown, Bull, Chue Hong, Filgueira Vicente, Grosser, Honda, A Jackson, Ji, Mai, A Smith, Spink, Steuwer** and **Weiland**.

How do we meet the computational demands of applications based on the bounds of circuit and communication technologies? How can we advance high-performance computing through innovation in languages, hardware, models, tools, and algorithms?

Since REF-2014, the Computer Systems theme has combined activities in the School of Informatics and EPCC, our translational research centre which aims to accelerate the effective exploitation of novel computing throughout industry, academia and commerce. EPCC is the largest group in the UK investigating how computing architectures enable world-class science and enhancing the understanding of high-performance computing (HPC) languages, algorithms and architectures. This has enabled growth in the richness and extent of research environment in this theme, with the CDT in Pervasive Parallelism (**O'Boyle**) providing increased research training, the establishment of new national research facilities such as ARCHER2 (**Parsons**), and our leadership of the UKRI-supported Software Sustainability Institute (**Chue Hong**), which facilitates better software to enable better research.

During the REF period, our research in this area has been recognised by 10 best paper awards and 2 test-of-time awards. Outstanding research includes **Nagarajan's** work on automatically generating directory cache coherence protocols (IEEE Micro Top Pick Honorable Mention, 2018), **Grot's** novel approach to balancing quality of service and throughput for colocated batch and latency-sensitive workloads (HPCA 2019 best paper), and using deep learning to optimise the allocation of resources in virtual 5G networks (**Patras**; MLN 2018 best paper).

Four of our submitted impact case studies arise from this theme – and three of them from research carried out during the current REF period: compiler fuzzing approaches for systematic testing of the Android runtime (**Franke, Leather**) used for every new release of Android; code size optimisation for WiFi chipsets (**Franke**) deployed at Qualcomm in over 1 billion devices; and

the NextGenIO persistent memory layer (**Weiland, A Jackson, Parsons**) incorporated into a product family by Fujitsu. The fourth impact case study is the low-power embedded processor design (**Topham**) commercialised by Synopsys and now used in over 2 billion devices.

B. Data Science and Artificial Intelligence (DS&AI)

Theme leaders: **Rovatsos, Simpson, Storkey**

45 category-A staff; £26.3M research awards; 105 current PhD students

26 new category-A staff in the REF period: **Albrecht, Belle, Chadwick, Court, Crowley, Gal, Goddard, Gutmann, Kökciyan, Luz, Ma, Mac Aodha, Manataki, Martin, Narayanaswamy, Nikolic, Onken, Oyarzún, Papapanagiotou, Restocchi, Seth, K Smith, Sorokina, Sun, Tsanas** and **L Webb**.

What are the computational foundations of data science and artificial intelligence? How can these be applied to thematically important applications such as biomedicine, health and wellbeing, software and algorithm engineering, and financial services?

Edinburgh has a long and influential history in DS&AI, and DS&AI is at the core of the university's strategic priorities. We have substantially extended the area with 20 new hires, particularly in **biomedical AI, machine learning**, and **reasoning** – areas in which we have increased research training through the **CDTs** in **Data Science** and **Biomedical AI**. In Biomedical AI, we have developed targeted interdisciplinary collaborations through the Usher Institute (Subsection 1.1.E), the MRC Institute for Genetics and Molecular Medicine, the Centre for Systems and Synthetic Biology, and EPCC's data safe havens for medical researchers. Academic impact in this area includes **Sanguinetti's** best paper award at the European Conference on Computational Biology 2016 for a probabilistic model to extract higher order methylation features for clustering and prediction in epigenomic studies, significant advances in the exploration of the stability of network activity in cultured rat hippocampal neurons (**Hennig**, J Neuroscience, 2015), proteomic analysis of postsynaptic proteins in regions of the human neocortex (**Sorokina, Armstrong**, Nature Neuroscience 2018), and gene expression analysis to dissect the molecular control of cellular processes in the human brain (**Simpson**, Nature Communications 2017).

Major findings in machine learning and probabilistic modelling include: **Sutton's** work on using machine learning approaches to model source code, which has helped to define this new field (ACM SIGSOFT Distinguished Paper Awards at FSE–2014 and ICSE–2020); **Storkey's** development of the “Neural Statistician” (ICLR 2015) introducing an innovative way of efficient learning from new datasets for both unsupervised and supervised tasks; **Murray's** influential work on deep generative modelling (ICML 2015, NeurIPS 2017); **Gal's** framework for computing fair solutions to rent division problems (EC 2016 best paper award); and a new computational

framework for inducing probabilistic logic programs over continuous data (**Belle**, ILP 2020 best student paper award).

We have developed new collaborations on using data science and AI techniques for science, recruiting a senior data scientist specifically focused on this area (**Seth**).

Impact from DS&AI includes: the analysis and modelling of laboratory animal behaviour (**Armstrong**, PLoS One 2017) commercialised by a startup (Actual Analytics) and in use at companies including AstraZeneca, GSK and Bayer; the use of speckled computing for improved air pollution and health monitoring (**Arvind**, Euromicro DSD 2016); **Williams'** work on the PASCAL Visual Object Classes Challenge (IJCV 2015) which had a massive effect on the fields of machine learning and computer vision.

C. Foundations of Computation

Theme leaders: **Kiayias, Libkin**

43 category-A staff; £29.2M research awards; 65 current PhD students

20 new category-A staff in the REF Period: **Campbell, Cao, Elahi, Garcia-Patron Sanchez, Guagliardo, Guo, Heunen, Kammar, Kapourniotis, Kiayias, Kohlweiss, O'Connor, Pieris, Polgreen, Ricciotti, Velichkov, Waites, Wallden, Zacharias** and **Zikas**.

What are the theories which underlie the analysis and design of computer systems? How can these foundational theories be implemented and applied practically?

Edinburgh's tradition in advancing the foundations of computation goes back five decades, with research in the area spearheaded by the Laboratory of the Foundations of Computer Science (LFCS), which focuses on the theoretical underpinning of computer science and its applications. The key objective of this theme is to understand the fundamentals of computing, providing the discipline with core scientific principles that can be used to guide computing practice and interaction with other sciences. Work in this theme encompasses a broad range of activity from blue-sky research, through collaboration with industry, to work with standards bodies. The real-world impact of our work is strengthened through large-scale joint academic-industry research with IOHK (**Kiayias**; £3.3M; Blockchain Technology Laboratory) and Huawei (**Fan, Hillston**; £7.6M; data management and processing).

In the REF period, we created a new group in **cyber security and privacy**, through a new chair (**Kiayias**) and four further academic appointments (**Elahi, Kohlweiss, Velichkov**, and **Zikas**), that is recognised as an Academic Centre of Excellence in Cyber Security Research by NCSC. Our fundamental research has had considerable impact: the Cardano ADA cryptocurrency, one of our impact case studies, is underpinned by theoretical work on the first provably-secure proof-of-stake protocol, Ouroboros (**Kiayias**, CRYPTO 2017).

Our work in databases continues to be highly influential. **Libkin** and **Guagliardo** have performed pioneering work in reasoning over incomplete information, including a new framework for answering queries over incomplete databases (best paper award KR2014) and sound logics for handling incomplete information (best paper award KR2018). **Fan** developed a new approach to parallelising sequential graph computations (best paper award ACM SIGMOD 2017) which has now been deployed at scale by Alibaba. **Nikolic** answered a deep open problem about the complexity of incremental maintenance of join queries (best paper award ICDT 2019).

During the REF period, we have strengthened **quantum computing**, through membership of the **Networked Quantum Information Technologies Hub**, with standout work on quantum verification (**Kashefi**, Phys Rev A 2017) and on categories for quantum theory (**Heunen**, **Kammar**, LICS 2017).

Fundamental research on programming languages continues to be highly influential, including effect handlers (**Plotkin**, and **Kammar**, J Functional Programming 2019), now a key abstraction in Uber's open-source Pyro probabilistic programming language and Github's internet-scale analysis of source code, and session types (**Wadler**, J Functional Programming 2014).

Other significant work includes the first polynomial-time algorithms for computing optimal reachability probabilities for Branching Markov Decision Processes (**Etessami**, Information and Computation 2018), and the RABIT/Reduce tool for Buchi automata minimization and language inclusion (**Mayr**, LMCS 2019).

D. Language, Interaction, and Robotics (LIR)

Theme leaders: **Lapata**, **Vijayakumar**, **B Webb**

49 category-A staff; £47.7M research awards; 254 current PhD students

29 new category-A in the REF Period: **Alexandru**, **Bach**, **Bell**, **Bilen**, **Birch-Mayne**, **Germann**, **Haddow**, **Heafield**, **Hill**, **Hospedales**, **Ivan**, **Khadem**, **Li**, **Lopez**, **Magdy**, **McNeill**, **Mistry**, **Mollica**, **Ribeiro**, **J Robertson**, **Ross**, **Sennrich**, **Sevilla-Lara**, **Subr**, **Titov**, **Tobin**, **Tonneau**, **Valentini** **Botinhao**, and **Vaniea**.

How can we build computational and mathematical models of language and cognition and link them to the real world? How can we use data to design tangible products and intangible services to transform the ways we work, live at home, care for each other, and play? How can we build robotic devices and autonomous systems that achieve human-like performance at complex tasks in real-world physical environments?

The underlying aim of the LIR theme is to devise computational methods which promote our understanding of how agents (humans or machines) interact with each other and the physical world, and to build systems which transform the ways humans interact with each other and with

machines. Key areas include language (written and spoken), vision, graphics, learning, robotics, and motor control.

During the REF period we have continued to invest strongly in **robotics**, with the establishment of the national **Robotarium** and the **CDT in Robotics and Autonomous Systems** (with Heriot Watt University), and in **natural language processing (NLP)**, with a new **CDT in NLP**.

A distinguishing feature of Edinburgh NLP research is the effective combination of sophisticated linguistic modelling with innovative probabilistic machine learning, and deep learning approaches to NLP. Research highlights from this substantial research grouping include large-scale semantic parsing without question-answer pairs (**Steedman, Lapata**, TACL 2014), and neural machine translation using byte pair-encoded subword units (**Sennrich, Haddow, Birch-Mayne**, ACL 2016) – highly cited work which underpins modern neural machine translation, and forms the basis of one of our impact case studies, with neural machine translation software developed at Edinburgh used by over 300 million people in Microsoft's products. **Lapata** (IEEE TPAMI 2017) pioneered work on the fusion of linguistic representations of word meaning together with visual representations to derive grounded meaning, which was the basis of her BCS Karen Spärck Jones lecture (2019).

Our work on speech processing underpins two of our impact case studies: automatic speech recognition commercialised by Quorate Technology and used by Hansard to construct the UK's official parliamentary record (**Bell, Renals**, IEEE TASLP 2016); and articulatory modelling and audio-driven animation commercialised by Speech Graphics and used by nine out of the ten top games publishers (**Shimodaira**, IEEE ICASSP 2014). **Goldwater** (Computer Speech and Language 2017) proposed a segmental framework for fully unsupervised large-vocabulary speech recognition, which featured in her BCS Roger Needham Award Lecture (2016).

Research highlights in robotics include humanoid balancing behaviour featured by underactuated foot motion (**Li**, IEEE Trans Robotics 2017), multimodal models for prosthetic hand control (**Vijayakumar**, J Neuroengineering and Rehabilitation 2017), and robots with insect brains (**B Webb**, Science 2020). Our work in robotics has had impact through robotic intelligent warehousing deployed by Hitachi (**Vijayakumar**), the startup Robotical (PhD graduate Enoch), and through our participation in three of the four EPSRC/ICSF Robotic Hubs.

Significant research advances in graphics and computer vision include phase-functioned neural networks for character control (**Komura**, ACM Trans Graphics 2017) and a deep learning approach to freehand sketch recognition which outperforms humans (**Hospedales**, IJCV 2017). In cognitive science, research highlights include modelling human reading with neural attention (**Keller**, EMNLP 2016) and investigating how easily people learn an unusual physical or social causal relation from a pattern of evidence (**Lucas**, PNAS 2017).

E. Interdisciplinary research structures

Informatics is at the collaborative heart of the university, and we have created an environment to nurture interdisciplinary partnerships and alliances, resulting in eight interdisciplinary research centres, three of which have been initiated during the REF period (discussed in more detail in Section 4); for each we name the lead researcher from Informatics:

- **Bayes Centre** (director: **Rovatsos**), the university's innovation hub for Data Science and AI, bringing together researchers from across the mathematical, computational, engineering, and natural sciences. The £45M Bayes Centre connects Informatics to the wider Data-Driven Innovation (DDI) programme and its other innovation hubs, specifically the Edinburgh Futures Institute (in Arts, Humanities, and Social Sciences) and the Usher Institute (in Medicine and Veterinary Medicine).
- **Centre for Design Informatics** (lead: **Wolters**, with a new chair (Vines) starting in January 2021), in collaboration with Edinburgh College of Art, with a focus on technological innovation through design with data and processes, e.g. approaches to building immersive data visualizations (**Bach**, IEEE TVCG 2018).
- **Centre for Research in Digital Education** (co-director: **J Robertson**), in collaboration with Moray House School of Education, researching how technology, culture, learning and policy intersect within research and practice in digital education, for instance using design theories to create cognitive training games (**J Robertson**, IJHCS 2019).
- **Centre for Speech Technology Research** (lead: **Renals**), in collaboration with Linguistics, concerned with all aspects of speech processing, e.g. using ultrasound sensing for child speech therapy (**Ribeiro, Renals**, Interspeech 2019).
- **Centre for Doctoral Training in Biomedical AI** (director: **Simpson**), established in 2019 in collaboration with Biological Science, Institute of Genetic and Molecular Medicine, Mathematics, Regenerative Medicine, Social Sciences, and the Usher Institute.
- **Edinburgh Futures Institute – EFI** (lead: **Magdy**), established in 2019, is the university's hub for DDI in the arts, humanities and social sciences, with five key research programmes: Ethics of Data and Artificial Intelligence, Critical Infrastructure, Creative Informatics, Financial Technology, and Data Civics and the future of public services. Interdisciplinary research with EFI includes social media analysis of Brexit (**Hill**, JCMS 2019); strength-in-places award GOFCoE (£23M 2020) to deliver data-driven advances in financial services data and fintech innovations (**Ma**).
- **EPCC** (director: **Parsons**), the university's supercomputing centre, with interdisciplinary collaborations spanning physics and astronomy to medicine. EPCC hosts and administers the Edinburgh International Data Facility, which underpins the university's

DDI programme. It won the contract to host and administer the ARCHER2 national supercomputer (£82M).

- **Usher Institute** (lead: **Luz**), an institute of the Medical School, with a focus on the data-enabled transformation of health, for instance the assessment of Alzheimer's disease using linguistic information (**Luz**, IEEE J Selected Topics in Signal Processing 2020).

Interdisciplinary impact includes the influential IUPHAR/BPS Guide to Pharmacology which draws on our data citation work (**Buneman**, Nucleic Acids Research 2016) and informing and influencing the design, development, and delivery of the Scottish school curriculum for computer science across 2,514 schools (**J Robertson**, CACM 2017).

1.3 Ethics, integrity, and open research

A. Research ethics

Given our vision to support social transformation, it is essential that our research outputs are developed to the highest ethical standards, especially given the ubiquitous and transparent nature of many of the applications of our research. Thus, responsible research and innovation has become core to research design in Informatics. For example, we were one of the first to invite the EPSRC-funded ORBIT team to present a Responsible Research and Innovation (RRI) workshop in June 2017 and in 2020 we hosted a bespoke Ethics Training Workshop for Informatics organised by the Association of Research Managers and Administrators.

As well as supporting this approach amongst staff, it is a key element of PGR training with each of the CDTs incorporating courses on RRI. In addition, in response to GDPR, we have developed a specific focus on clarification of the legal basis for processing data for research, provision of a template data protection impact assessment when required, and clarification of participant rights in relation to their data. All staff and PGR students are required to take information security and data protection training.

Commitment to maintaining the highest standards of research integrity is supported by the university. Edinburgh was one of the first universities to adopt the UKRIO Code of Practice for Research, which underpins the UUK concordat resulting in a Research Ethics and Integrity Framework specific for STEM disciplines (REF5a). This framework includes ethics review procedures and protocols, an implementation of the university's research misconduct policy and training in ethics and integrity.

In line with this framework, we have renewed the School of Informatics research ethics procedure, seeking to ensure full legal and ethical accountability, within a process that ensures that decision-making is as close to the pertinent research expertise as possible. Similar procedures are in place for EPCC, which hosts the Scottish National Data Safe Haven on behalf

of the Scottish Government and NHS Scotland, and the Bayes Centre – our researchers in the Usher Institute follow the best-practice ethical review procedures in the College of Medicine and Veterinary Medicine.

Ethics and responsible computing drives research across all four themes, for example issues of fairness in AI systems, and underpins our new research programme in AI and Social Good (Section 1.5). In 2019 the university appointed Vallor (UoA-30) to the £5M-endowed Baillie Gifford chair in Ethics of Data and AI, and Informatics is actively participating in the newly formed Centre for Technomoral Futures in EFI.

B. Open and reproducible research

We have substantially increased our support for open and reproducible research. The Software Sustainability Institute (director: **Chue Hong**) is based on the notion that good software engineering practices lead to better software, which in turn makes research more reusable and reproducible. To this end, we have developed the notion of Research Software Engineers as key to reproducibility, established a series of software, data, and library carpentry programmes, and prepared guides and online evaluations for software sustainability.

EPCC shows leadership in this field through a number of funded projects specifically concerned with open data. This includes the European Open Science Cloud Hub which has developed a compute and storage platform, Health Data Research UK which supports the use of shared data for public health research, the National Safe Data Haven providing research access to routinely collected medical images, and the Scottish Administrative Data Research partnership which supports research to develop policy based on the use of shared and linked data.

Our approach to research publications is largely based on a Green Open Access (OA) policy, although financial support is provided for Gold OA. Since 2016, 97% of our publications have complied with the REF OA policy, and many of our research outputs are made available through arXiv, in addition to being published on the Edinburgh Research Explorer repository. We host regular workshops for all researchers and administrators on copyright, OA, and publishing. Our researchers have provided leadership for several major open access publication venues, e.g. editor-in-chief of Journal of ORS (**Chue Hong**), managing editor of Logical Methods in CS (**Plotkin**), programme chairs of EACL (**Goldwater**, **Lapata**) and ICLR (**Murray**), publications chairs of ACL (**Cohen**) and ICML (**Murray**), senior editor of IEEE Open Journal of Signal Processing (**Renals**), editor of Quantum (**Heunen**), and advisory board member of NeurIPS (**Williams**).

Open-source software enables reproducible research: we have led several open-source software projects including the Marian machine translation system (**Birch-Mayne**, **Haddow**), the Met Office NERC Cloud model (MONC; **Brown**, **Weiland**), and the DexFuzz testing framework for the Android Run Time system (**Franke**, **Leather**). Other open-source projects which we lead

include VirtualFlyBrain (**Armstrong, Court**), an interactive, web-based tool to explore the detailed neuroanatomy, transgene expression, and associated phenotypes of the fly's brain, Hapbin (**Chue Hong**) a tool for haplotype-based scans for detecting signatures of selection in Large Genomic Datasets, and NetMap a framework for fast packet I/O (**Honda**).

1.4 Impact strategy

To maximise real-world impact arising from our research, we have developed an impact strategy based on five routes to deliver economic and social impact:

- **Collaboration with large companies** – exemplified by five impact case studies in which our research is deployed in products developed by Facebook, Fujitsu, GitHub, Huawei, Microsoft, Qualcomm, Synopsys, and Uber;
- **Collaboration with SMEs** – exemplified by three impact case studies in which our research has delivered impact through products and services deployed by IOHK, Lingo24, and Emotech;
- **Formation of start-up and spinout companies** – exemplified by two impact case studies in which our research was commercialised by the Edinburgh start-up companies Speech Graphics and Quorate Technology;
- **Collaboration with public / third sector organisations** – e.g., our work delivered impact through the UK Meteorological Office and multiple UN agencies such as WIPO (the World Intellectual Property Organisation);
- **Release as open-source software** – exemplified by three impact case studies based on the Android Run Time System, the Marian machine translation software, and the Pyro PPL.

Evidence for the success of this strategy is given through our impact case studies, and additional impacts discussed in Section 4.2. Informatics is at the centre of a rich ecosystem of SMEs in Edinburgh, e.g. we are ten minutes' walk from CodeBase, which houses over 100 technology SMEs. This proximity and frequent interaction provides substantial mutual benefit.

During the REF period we have invested to increase the effectiveness of our pathways to impact, with specific emphases on:

- Support for commercialisation by the **Informatics Commercialisation Team**;
- Translational research and development at **EPCC**;
- The Edinburgh City Deal **DDI** programme and the **Bayes Centre**;
- Establishment of joint **industry-university translational research centres**;
- **Innovation and industrial fellowships**;
- **Outreach and public engagement**.

The **Informatics Commercialisation Team** (director: Ashe) was established by Adams (retired 2016, CBE 2017 for innovation and entrepreneurship), and is recognised for catalysing the entrepreneurial digital culture in Edinburgh, demonstrated by Skyscanner (acquired for US\$1.75B in 2016), and technology incubators such as Codebase and Techcube. During the REF period we have invested £5M to refurbish space for startup and spinout companies in Appleton Tower. Start-ups are identified and supported through a variety of mechanisms (detailed in Section 3.4). Over 50 successful start-ups and spinouts have been established from our research in Informatics over the past decade (23 in the current REF period), including Speech Graphics, Quorate Technology, Actual Analytics, Imandra, Aveni, and Robotical.

During the REF period we focused our business development infrastructure on increasing industrially supported research, resulting in three of our five joint **industry-university research centres**: the **IOHK** Blockchain Technology lab (**Kiayias**; £3.3M), the Edinburgh **Huawei** Research Lab (**Fan, Hillston**; £7.6M), and the **ARM** Centre of Excellence (**O'Boyle**; £300k). We have also established funded research collaborations with over 30 companies, several of which are reflected in our impact case studies. These activities cover all four of our research themes, encompassing funded collaborations with companies such as Oculus VR / Facebook (£787k), Samsung (£1M), and Microsoft (£666k). Other such collaborations, including co-supervised PhDs, include Amazon, Bloomberg, FiveAI, Google, Hitachi, and Toyota.

EPCC, the university's supercomputing centre, and the only such national-scale facility hosted in a university in the UK, is central to our impact strategy through its provision of world-class computing facilities for scientists across a range of areas – from particle physics to cosmology, novel materials to combustion, and climate modelling to energy – and generates new techniques and approaches that are taken up by industry. EPCC houses and administers an exceptional range of supercomputers – including ARCHER, ARCHER2, CIRRRUS, DiRAC, RDF, and EIDF with a total contract value of over £180M, with 104 staff – mainly software architects, software engineers, and programmers – providing companies of all sizes with access to novel computing expertise and specialist computing facilities (Sections 3.3 and 3.4).

A key aim of EPCC is the provision of HPC to companies, in particular SMEs, via the Accelerator on-demand HPC service. Examples include aerodynamic modelling for high-performance sports car manufacturer Koenigsegg, and development of very large-scale predictive analytics for Global Surface Intelligence (GSi).

The Edinburgh City Deal's **DDI** Programme, established in 2018, is key to our future impact strategy with ambitious targets to train 100,000 people in data science in the local region and to support 1,000 local companies to make better use of data in their businesses. The **Bayes Centre** is the DDI hub, hosting 38 external organisations across private, public and third sectors, 28 of which are startups, spinouts or SMEs with their principal R&D activities at Bayes. Four

major corporates have their industry/university research centres at Bayes: the **Royal Bank of Scotland** Data Innovation Unit, the EPSRC Prosperity Partnership ASiMoV with **Rolls-Royce** (**Parsons**; £8.1M), **SAS Data Analytics** (£2.6M in software licenses), and the Edinburgh **Huawei** Research Lab. Furthermore, Informatics Ventures delivers a series of start-up and scale-up accelerators in the Bayes Centre. In particular the **Wayra AI and Blockchain Accelerator** (funded £567k in 2018 by Telefonica), is developing an early-stage start-up programme targeting entrepreneurial PhDs and early career researchers.

Five of our researchers have been awarded industrial or innovation fellowships: **Birch-Mayne** (UKRI Innovation Fellowship, robust neural machine translation); **K Smith** (UKRI Innovation / Rutherford Fund Fellowship, network science for human brain structure and function and protein-protein interaction networks); **Albrecht** (RS Industry Fellowship with FiveAI, AI technologies for autonomous driving in urban environments); and **Komura** (RS Industry Fellowship with Disney, natural and realistic movement of interacting animated characters). **Kohlweiss** held a three-year industry fellowship supported by IOHK.

Our work on **outreach and public engagement** (Section 4.2) has developed substantially in recent years, and we have a communications team of seven people spread across the School of Informatics, EPCC, and the Bayes Centre. Examples of our public engagement in the REF period include **Vijayakumar's** TedX and New Scientist Live talks, the "Robots" exhibition at the National Museum of Scotland in 2019 (over 60,000 visitors) and Harmonium (Oberlander†, **Hill**) which opened the 2015 Edinburgh International Festival with a live audience of 20,000 people.

1.5 Future strategic aims and objectives

To bring about our vision of a transformed digital society, we continue to focus on foundational research which can be brought to bear on societal problems through collaboration with academic and industrial partners. To achieve our envisaged societal impact, we are planning the development of three new research areas, combining fundamental curiosity-driven research with the potential for substantial real-world impact:

- **AI for Social Good:** Led by **Hillston**, this new research programme aims to increase our activity around the beneficial use of AI, and associated ethical issues, with the objective of weaving the ideas of AI for social good through all our AI activities. Working closely with the EFI and the Centre for Technomoral Futures, and aligning with the university's focus on the UN Sustainability Goals, our aim is to develop an interdisciplinary CDT, and to establish a research visitor programme centred on social challenges and sustainability, working closely with related national initiatives, such as the Alan Turing Institute programme on Data Science for Social Good.

- **Beyond Moore's Law:** Led by *O'Boyle* and *Parsons*, this programme will look afresh at the entire computing eco-system, from applications to technology, to tackle the challenge of computing growth now that Moore's Law is dead. This programme will draw researchers across all our research themes building on our strength in computer architecture, security, optimizing compiler technology, the science of computation, supercomputing, and programming language innovation, as well as drawing in researchers from deep learning and robotics.
- **TEAMER: TEAching MachinEs to Reason Like Humans:** Led by *Lapata*, we are establishing a new research centre to unify multiple strands of AI research to develop AI tools with reasoning abilities which go beyond the pattern matching approaches which characterise the state-of-the-art. This centre will bring together areas of existing strength including deep learning, probabilistic programming, statistical knowledge bases, and explainable AI, and foster greater collaboration with the ATI, the School of Philosophy, Psychology and Language Sciences, and Design Informatics.

2. People

2.1 Staffing strategy and staff development

In order to achieve our vision of a transformed society fit for the digital age, our staffing strategy is primarily driven by attracting, supporting, and empowering the best researchers. To attract internationally leading researchers, we ensure that our environment and infrastructure are stimulating, flexible, and supportive. We support researchers through tailored mentoring and training throughout their career, but especially early in their career through a significantly decreased teaching and administration load. We foster a strong sense of community and collegiality, by offering researchers affiliation to one or more research institutes that form their academic family. Researchers are empowered through excellent professional support services ranging from help with grant writing through to advice on IP generation and entrepreneurship.

Following REF-2014, we embarked upon a strategy of planned growth whilst maintaining excellence. We prioritised investment in academic staff and in scholarships for postgraduate research students. We have largely recruited at the lecturer (grade UE08) or reader (grade UE09) levels, with external recruitment to chairs (UE10) exceptionally undertaken to provide strategic leadership in new or under-developed areas of research and curriculum. During the REF period we have externally recruited three new chairs – **Kiayias** (Cyber Security and Privacy), **Martin** (Computer Science), and **J Robertson** (Digital Learning) – with a further twelve of our staff promoted to personal chairs: **Aspinall** (Software Safety and Security), **Cole** (Patterned Parallel Computing), **Etessami** (Algorithms, Games, Logic and Complexity), **Goldwater** (Computational Language Learning), **Kashefi** (Quantum Computing), **Komura** (Computer Graphics), **Marina** (Networked Systems), **Pain** (Interactive Learning Environments), **Rovatsos** (Artificial Intelligence), **Sanguinetti** (Computational Bioinformatics), **Stevens** (Mathematics of Software Engineering), **Storkey** (Machine Learning & Artificial Intelligence). An externally recruited chair in Design Informatics (Vines) is starting in January 2021.

Currently, Informatics at Edinburgh comprises 164 category-A researchers (144.14 FTE), a further 102 research staff, 8 university teachers, 85 software architects / engineers / programmers, 38 computing / technical support staff, 101 administrative staff, and 504 postgraduate research (PGR) students. Compared to REF-2014 this represents 52% more category-A researchers (by FTE), and 80% more PGR students.

As discussed in Section 1, in the REF period we have appointed 64 new independent researchers and enabled a further 28 existing staff to become established as independent researchers. This has resulted in a total of 92 new independent researchers included in this UoA, who were not returned in the previous REF. We now include 7 category-A staff from EPCC.

During the same period, 31 of our research independent staff have left this UoA, including 7 retirements. Many leavers have secured leading industry positions at organisations such as Google (Ferrari, Viglas), Bloomberg (Osborne), Salesforce (Lee), and Anaconda (Bednar), or have taken full professorships at leading institutions (Diakonikolas, Wisconsin; Koehn, Johns Hopkins; Maneth, Bremen; Sanathanam, Oxford; Yamagishi, NII Japan).

The net result is a young vibrant community and an increase of 60 independent researchers (49.29 FTE), compared with REF-2014.

During the REF period the university has significantly invested in **Chancellor's Fellowships**, four of whom were appointed to this UoA (**Belle**, **Cao**, **Papapanagiotou**, and **Tsanas**), with **Belle** subsequently winning a Royal Society URF. The Chancellor's Fellowship scheme provides high potential researchers with a supported five-year, tenure-track, personal development programme, allowing them to primarily focus on their research portfolio, with support including dedicated training and networking events, gradually taking on wider teaching and academic management duties.

Staff who leave are replaced, although not necessarily on a like-for-like basis as the opportunity is taken to review needs and opportunities available in the light of a fast-developing research and teaching landscape. Similarly, where staff secure substantial (five-year) fellowships these are backfilled with a new appointment to an open-ended post, on the basis that these latter appointments will be making a substantive contribution to the School's research and teaching at the end of the five-year period.

We recruit internationally, aiming to attract the best. We have thus developed a strongly diverse body of staff. Category-A staff come from 31 countries, with 36% from the UK, 29% from the EU and EEA, and 35% from the rest of the world.

Our strategy has led to a staffing profile that is balanced by role (Table 1). The biggest increase is in the number of research fellows, primarily due to the development of staff into independent researchers. (All Chancellor's Fellows at REF-2014 progressed to lecturer/reader positions.)

Staff type (head count)	REF-2014		REF-2021	
Professor	39	38%	45	28%
Reader / Senior Lecturer	24	23%	45	28%
Lecturer	17	16%	40	25%
Chancellor's Fellow (tenure track)	10	10%	4	2%
Research Fellow	14	13%	30	18%
TOTAL	104		164	

Table 1: Category-A staff in REF-2014 and REF-2021.

In addition, in the REF period, seven post-doctoral researchers were successful in securing open-ended Lecturer and Reader appointments through open competition (**Bell, Crowley, Guagliardo, Hill, Sennrich, Wallden, Wolters**).

Staff are encouraged and supported to apply for **externally funded fellowships**. It is our aim, at any time, to have 10–12% of research and teaching staff on fellowships. During the REF period, 23 of our staff were supported by external fellowships (*EPSRC: **Birch-Mayne, Heunen, Kashefi, Libkin, Martin, O'Boyle, Sun**; Royal Society: **Albrecht, Belle, Kammar, Komura, Subr**; ERC: **Cheney, Fan**, Ferrari, **Lapata**, Santhanam, **Steedman** (twice), **Titov**; Royal Academy of Engineering: **Leather**, Petoumenos; UKRI: **Khadem, K Smith**), with 16 currently supported in this way.*

Staff are supported to spend time in **industry** in order to assist in their career development and to foster strong industry-academia links. Arrangements include periods of full or partial leave-of-absence, e.g.: **Bhatotia** (Huawei); **Hospedales** (Samsung); **Lopez** (Rasa); **Murray** (Amazon); **Plotkin** (Google; Microsoft); **Ramamoorthy** (FiveAI). Also, four people whose substantive employment is in industry hold part-time positions within the School and contribute to building links between industry and academia, development of research strategy and curriculum: Bishop, Gordon, **A Smith** (Microsoft Research), and **Sutton** (Google).

Many of our staff perform commercial consultancy, e.g., **Bell** (Quorate), **Cohen** (Huawei), **Heafield** (Amazon; Intel), **Kiayias, Wadler**, and **Zikas** (IOHK), and **Topham** (Synopsis), enabling the translation of world-class research knowledge into tangible impact for our commercial partners.

We place a strong emphasis on internal promotion. There have been 68 promotions for our category-A staff during the REF period (Table 2), with promotion decisions based on contribution across research, impact, teaching, and contribution to administration and leadership.

Year	2013/14	2014/15	2015/16	2016/17	2017/18	2018/19	2019/20
To Senior Researcher (UE08)	1	1	1	1	3	6	7
To Senior Researcher (UE09)	0	1	0	0	0	1	4
To Senior Lecturer / Reader (UE09)	6	5	4	4	3	2	6
To Chair (UE10)	1	2	0	1	2	3	3

Table 2: Promotions of category-A staff during the REF period.

Annual workshops help staff identify their stage of readiness for **promotion**, the actions that they need to take, and the processes involved in applying.

Annual **Performance and Development Review** (PDR) is central to our staff development process. Training sessions are provided, for both appraisees and appraisers, on how to obtain the most from the PDR process. We focus on training for new PIs as they take on responsibility for management of a research team, and to develop the next cohort of senior academic managers and leaders. The university operates annual leadership and senior leadership programmes. We participate in these and are developing specific Informatics programmes (27 staff will benefit during AY20/21). We also nominate participants for inclusion in the university's allocation of places on the Advance HE Aurora leadership development initiative for women – in the REF period five of our staff have participated.

A staff development budget enables staff to apply for funding to attend external courses or conferences, covering fees, travel, subsistence and additional child-care costs, as well as time allowance and/or time off for study. We have invested over £65k in staff development in the REF period.

We also create customised local courses, for instance training in recruitment law, appraisal, unconscious bias (in addition to the university's online provision), dealing with students with specific learning profiles and disabilities, harassment and bullying, and applying for fellowships.

Our generous **sabbatical** policy for research and teaching staff aims for 8% of research and teaching staff to be on sabbatical each year. Entitlement to a sabbatical accrues at one semester for each four years of service, with most staff taking a full year sabbatical after eight years or thereabouts. In this period 44 staff have used sabbaticals to advance their research, impact or teaching, or a combination of these.

Support for early career researchers: Early career researchers (ECRs) comprise post-doctoral research associates (PDRAs; grade UE07) and senior research associates (UE08), and staff appointed as lecturers or research fellows (UE08), including those who may be a co-Investigator on research grants but do not yet hold a research grant as a principal investigator.

ECRs benefit from a supportive research community structure within Informatics and their respective research groups. This involves seminar series, formal and informal mentoring, and advice and workshops on applying for research funding. Guided by the Concordat to Support the Career Development of Researchers, we support researchers at each stage of their career, with targeted support for PDRAs and for Lecturers/Research Fellows. A workshop for Informatics and EPCC staff on the Concordat was held in January 2020, and the university's Institute for Academic Development (IAD) offers a range of relevant courses in the areas of obtaining research funding, personal effectiveness, research governance, and engagement, influence, and impact.

Lecturers and Research Fellows are supported in their career development towards becoming research leaders, with relief from teaching and administrative duties in their first year to enable

them to establish themselves as part of the research community. In addition, they are provided with a package of start-up funding tailored to their needs, including fully-funded PhD studentships, equipment, travel funds and other resources.

We provide informatics-specific training for PhD supervision, jointly with the IAD, and early career researchers are paired with a more experienced research training supervisor.

Workshops are provided on preparing grant applications, including applications for fellowships. Lecturers and Research Fellows are provided with a mentor to help guide them through the process and grant applications are internally peer reviewed, and constructive feedback is provided, prior to submission. Where appropriate, we also organise mock panels as part of their preparation.

Our **PDRA** population is 30% female, 70% male, and joins us from around the world. 27% are from within the UK, 39% from the EU and EEA, and 34% from the rest of the world. Currently 53% of our PDRA population are white and 40% BME and for 7% no data is available. PDRAs are managed by their research group leaders, who ensure (together with senior research managers) that they are engaged in the wider research community and that their longer-term career development is addressed. An explicit part of the research grant application process is to ensure that financial provision is made for travel, conference attendance, and collaboration opportunities for PDRAs included in the grant application.

The PDRA Development Co-ordinator (**Gutmann**) has specific responsibility to work with PDRAs and their managers to identify shared development needs and to identify or put in place opportunities to address these needs, e.g. through training courses arranged by the university's IAD or through in-house training and workshops. This post sits within a university-wide network of 'PDRA Champions' and support staff, coordinated by IAD. Our focus on career development for PDRAs is reflected in our increased success in promotion to Senior Researcher positions.

The Informatics Research Staff Society provides a forum for scientific, professional, and social interaction, for PDRAs across the research themes. Members of this society sit on the two major Informatics committees (the Senior Management (Strategy) committee and Research committee). A recently (2019) instigated lunch, held each semester with the head of Informatics, provides PDRAs with a forum to learn of developments within the university and to provide feedback and suggestions. Within research themes, we provide mentoring and advice for PDRAs applying for lectureship positions both at Edinburgh and elsewhere.

Staff recruitment: With the exception of 'named on grant' research roles, all researcher posts are recruited through open competition. Informatics follows university policies and procedures to ensure that all recruitment is transparent and fair. All selection panel members are required to successfully complete the university's on-line equality and diversity and unconscious bias

training, and selection panel chairs are required to also complete the university's Recruitment and Selection training.

Staff induction is the responsibility of line managers. New research staff are provided with a 'buddy', from within their immediate research group, who will guide them through their first few days and weeks at the university, to help them familiarise themselves with their new colleagues, working environment and relevant policies and procedures. In addition, new staff are provided with a senior staff member as a mentor who will support them in their longer-term personal development. As part of the induction process, new staff are provided with the opportunity to give feedback on their experience, both to help to identify any specific issues that require immediate remedy and to inform future review of the induction process.

2.2 Research students

We currently have 504 postgraduate research (PGR) students from 65 countries. Our PGR student population is 23% from the UK, 42% from the EU, and 35% from outside the EU, and is 21% female and 78% male. The gender balance is gradually improving: the intake for 2019–20 was 24% female, compared to 21% in 2014, and 14% in 2010.

Our research degrees include PhD, MPhil, and MSc by Research (MScR). The degree structure for CDTs either comprise a 1-year MScR programme followed by a 3-year PhD programme, or a 4-year PhD with integrated studies. Funding is provided from a wide variety of sources:

- EPSRC/UKRI CDTs (Data Science, Pervasive Parallelism, Robotics and Autonomous Systems, Natural Language Processing, Biomedical AI);
- EPSRC Doctoral Training Programme (DTP);
- EPSRC ICASE studentships;
- Industry sponsorship, e.g. ARM, Bloomberg, Huawei, Samsung, Toshiba;
- Externally funded scholarships, e.g. the Choksi and the Langmuir scholarships;
- University of Edinburgh scholarships;
- International competitive scholarships: Google PhD Fellowships, IBM PhD Fellowships, Microsoft PhD Scholarships.

In addition, we provide internal investment for PhD studentships, currently committing £2.5M per year, through internally-funded CDT places, pump-priming scholarships allocated to early-career lecturers, and competitive studentships and partially matched funding (e.g., in conjunction with industry funding).

All research students are Graduate School members. Each student is assigned a dedicated desk and a dedicated desktop PC, and has access to our research facilities, described in Section 3.3. We have a policy that any student with an accepted conference paper is provided

with the necessary funding for travel, subsistence and conference fees, with CDT students allocated a specific budget for this purpose.

All research students have a primary and secondary supervisor, responsible for guiding the student's research agenda, training them to become independent researchers, and monitoring their progress. All students receive training in responsible research and innovation. Each student's performance is measured against a sequence of specific milestones, with one milestone every 3-4 months throughout the PhD programme. An annual review makes recommendations about the student's progression. Our 4-year completion rate is currently 77%.

Our recruitment strategy is focused on **academic excellence** and **equality of opportunity**. We receive around 850 PGR applications per year; about 10% are selected, on academic merit. Applications have increased by 79% since REF-2014. Over 90% of our PGR students receive full funding (fees + stipend) for their studies. Admissions are needs-blind. We proactively monitor the gender balance of our intake throughout the recruitment cycle. When selecting for scholarships, if there are equally qualified candidates, we have a policy of prioritising those from under-represented groups.

PhD intake, the number of PhD completions, and school postgraduate research funding are shown in Table 3 and Figure 2. The Graduate School has grown by 80% since REF-2014, with 504 postgraduate research students, compared with 280 on the REF-2014 census date. In the REF period, we graduated 328 PhD students, an average of 47 per year, compared with the 195 PhD graduates (39 per year) that we recorded for REF-2014, an increase of 20% in the average number of PhD graduates per year.

Year	2013/14	2014/15	2015/16	2016/17	2017/18	2018/19	2019/20	Total
PhD intake:								
CDT/DTC	13	11	28	35	35	38	78	238
Other	48	44	36	50	63	66	75	382
Total	61	55	64	85	98	104	153	620
PhD completions:								
CDT/DTC	5	5	10	7	17	9	15	68
Other	30	38	44	33	39	33	43	260
Total	35	43	54	40	56	42	58	328
School funding (000s):								
Fees/£	238k	448k	403k	488k	622k	867k	1,209k	4,275k
Stipends/£	404k	591k	662k	776k	817k	1,099k	1,467k	5,816k

Table 3: PhD student intake, completions, and additional school funding

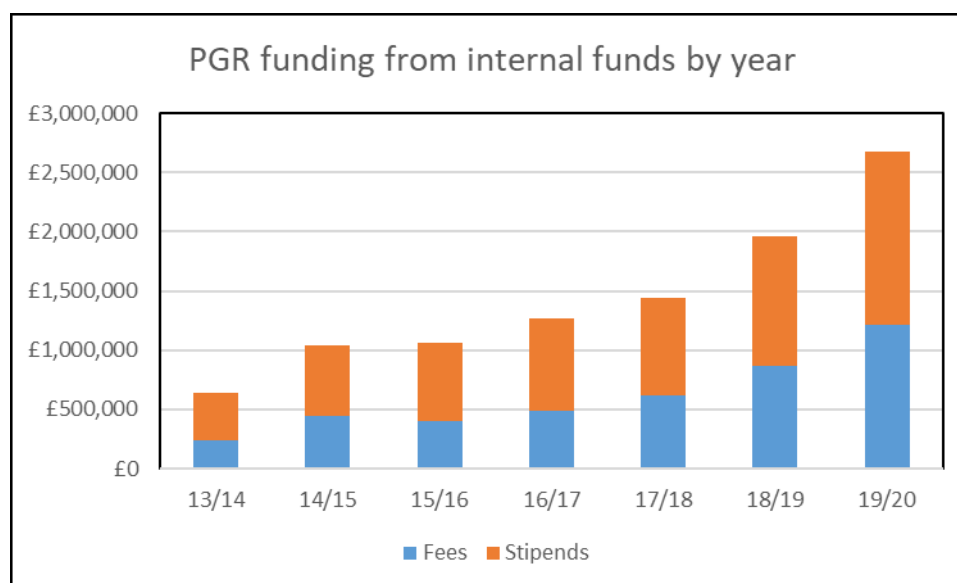


Figure 2: Internal postgraduate research funding (2013–2020).

Our PhD training courses address a wide variety of needs, for instance: experimental design, statistics, time management, research planning, proposal writing, viva preparation, oral presentations, networking, career planning, and obtaining and benefitting from internships. Most of these courses are designed and delivered in collaboration with the IAD, supplemented by two delivered by an external provider. Research students also have access to broader training opportunities provided by IAD, e.g. in public engagement. Equally importantly there is extensive training for supervisors also provided by IAD.

Dedicated workshops and training are provided by the Informatics business development team on commercialisation and intellectual property management.

To enhance the student experience and to foster peer support and social cohesion, the Graduate School financially sponsors a wide variety of social and academic events, which are proposed and organised by research students. These include a monthly board games night, a bi-weekly arts workshop, and poster events and workshops organised by the student-led society EdIntelligence.

Based on graduate surveys conducted within the REF-2021 period, an average of 98% of graduating PGR cohorts go into highly skilled employment or further study. An average 34% of each graduating cohort moves to a role in academia.

2.3 Equality, diversity, and inclusion

Equality, diversity, and inclusion (EDI) is of particular importance in Informatics, as we have a long history of working to address a very imbalanced gender ratio. Currently 18% of our category-A staff are female (compared to 17% in REF-2014); this ranges from 29% in the Language, Interaction, and Robotics theme to 7% in the Computer Systems theme. The proportions of female students are 24% (undergraduate), 32% (taught postgraduate), and 21% (research postgraduate – ranging from about 12% for the Computer Systems theme to 27% for the Language, Interaction, and Robotics theme).

Our commitment to gender EDI in Informatics is recognised by an Athena SWAN Silver Award, which was renewed during the REF period (2016). This award recognises – and encourages – better working practices, as well as measuring our progress. Our Athena SWAN action plan goes beyond gender equality and is focused on building an inclusive and welcoming community within Informatics, and which addresses all aspects of equality, diversity, and inclusion.

Informatics operates within the framework of the university's policies relating to EDI, dignity and respect, and family-friendly working. The latter have recently been reviewed and benchmarked against other universities within the Russell Group to ensure that they reflect best practice. All processes within Informatics, and the university, operate on the principles of equality, fairness, openness, and transparency. All new or revised policies seeking approval are required to identify any impact on EDI. The School of Informatics has strong commitment to these values with two senior members of the School (**Hillston, O'Boyle**) serving as Dignity and Respect Advisors in the university, and representation on the university Staff Pride Network (**Court, Dow**).

Flexible working: In order to accommodate flexible working patterns and family and caring commitments, Informatics buildings have unrestricted access to staff and students. Within the requirements of their research and any teaching and similar commitments, researchers can take advantage of remote access to computing systems, and this has worked very smoothly during

the Covid-19 lockdown. All committee meetings and general meetings are scheduled to be in the core hours of 10am-4pm, Tuesday to Thursday, and about 85% of all our meetings take place during the hours of 10am-4pm (based on analysis of room booking data). Where required, funding is available for additional childcare needs arising from conference attendance and similar events, assessed on a needs basis. Additional travel costs for staff with special needs are funded on the same basis.

Adjustments in terms of working hours, working arrangements, special equipment and support are made for staff with additional needs, usually on the basis of an assessment by the university's Occupational Health unit. Arrangements can be on a temporary or long-term basis, with longer-term arrangements reviewed annually. During the Covid-19 lockdown, and subsequently, flexible working or temporary reduction adjustments were made for those who requested them due to additional caring or childcare responsibilities.

The allocation of teaching and administrative duties is organised in a transparent work allocation model, aiming to share the duties fairly across staff. Adjustments are made for research activity, new staff, staff returning from absence, and staff with caring responsibilities.

Diversity: One of the attractions of working and studying within Informatics is its scale and diversity. We are one of the most ethnically, linguistically, and culturally diverse parts of the University of Edinburgh. Currently 67% of our category-A staff are white, 20% BME and for 13% no ethnicity data is available. They come from 31 countries; including postdoctoral research associates and research students, our researchers come from 103 different countries. We celebrate this diversity through culturally-focused events and recognition of celebration days, such as Eid, Nowruz, and Chinese New Year. We proactively seek to encourage applicants both internally and externally, from under-represented groups, and take advantage of the opportunities offered by some funders allowing additional applicants from such groups (e.g. the RAEng Fellowship scheme).

Two of our staff are Suffrage Science award winners (**Hillston, Martin**). We have increased the percentage of female members of the professoriate since the 2014 REF (13% to 21%). Since REF 2014, Informatics has had two female Heads of School – **Moore** (2014-2018) and **Hillston** (2018-date). **Hillston** also served in two previous senior leadership roles and three female academics, **Lapata, Lascarides** and **B Webb**, are currently in such roles. There are currently three BAME senior research leaders: **Fleuriot, Marina** and **Vijayakumar**.

EDI governance and networks: The EDI committee, which represents all staff and students (with representation from academic staff, professional staff, PDRAs, and PGR students), is directed by a senior academic (**Nagarajan**) who has a seat on the Informatics senior management committee. In addition, the Health, Safety & Wellbeing committee, chaired by the head of Informatics, considers issues related to disability and general well-being. We seek to

ensure diversity in academic appointment panels, aiming for overall gender balance and guaranteeing at least one female member on each panel. We also ensure male representation on committees (such as EDI) where women have traditionally been overrepresented. We participate in the university initiative to review EDI issues (gender, BAME and sexuality) in our curricula. For example, a working group on decolonising the curriculum is running a series of workshops with teaching staff and students to raise awareness and instigate a move to a more inclusive curriculum.

Informatics general meetings are held three times each year as an opportunity to provide updates on developments and obtain feedback on issues and concerns. During Covid-19 this has been increased to every 8 weeks, using a digital platform. A monthly newsletter and weekly email bulletins provide more frequent communications. A confidential 'suggestion scheme' provides a channel for anonymous feedback directly to the head of Informatics.

We partner in the College's Sabbatical Programme for Women that enables women close to applying for promotion to take a semester sabbatical to enhance their CV. Opportunities to attend specialist activities are funded (Women & Equality in the Workplace, BCS Women Lovelace Colloquium) and we organise events for celebration days (Ada Lovelace Day, International Day of Women and Girls in Science, International Women's Day, Trans Visibility Day). The School supports Hoppers (UG, PGT, PGR, researcher, academic women computing society), which celebrated its 15th birthday in March 2020.

To help develop an inclusive, cross-Informatics, community, we support and fund a range of extra-curricular social and sporting activities. Many of these are led by staff or research students, such as table tennis, French cinema club, arts and crafts, and football. We also have regular community-wide social events, such as the annual Informatics Jamboree and Ceilidh, as well as weekly yoga classes. In addition, research groups, and CDTs organise their own social and recreational events. The Staff Pride Network is active in promoting inclusivity through social events and other activities such as coffee and cake gatherings for LGBT colleagues and their allies.

Staff training: Training courses on Equality and Diversity Essentials and Challenging Unconscious Bias are now essential for all staff, and a pre-requisite for those on senior management committees or responsible for selection and recruitment. In addition, the staff induction process introduces new staff to the university's web-based E&D resources.

Equality, diversity, and inclusion in the REF-2021 submission: All staff involved with REF-2021 completed a specific training course on EDI for REF. We have followed policies and processes set out in our institutional code of practice in preparing the REF-2021 UoA-11 submission. Inclusion in the output portfolio is based solely on merit, as assessed by a panel of senior academics. 30% of our category-A FTE staff numbers are ECR and they contribute 26%

of REF2 outputs whilst 17% of our category-A FTE staff numbers are female and they contribute 15% of our REF2 outputs. Where appropriate, special circumstances have been considered in determining an individual member of staff's contribution to the output portfolio.

3. Income, infrastructure and facilities

3.1 Research income

Our research income (award value), summarised in Table 4, has grown significantly during the REF period to £167M compared to £63M in the REF-2014 period – over £1M per category-A staff member across the REF period. This is nearly double the average annual research income recorded in REF-2014: £23.9M/year (REF-2021) versus £12.7M/year (REF-2014).

	2013/14	2014/15	2015/16	2016/17	2017/18	2018/19	2019/20
Research Income (£/million)	23.5	16.7	18.7	20.2	24.2	34.1	30.1
Research Training Income (£/million)	-	8.6	-	-	-	-	13.5

Table 4: Research income by year

Research income comes from diverse sources (Table 5), with the majority coming from UKRI (£69.0M) and EU H2020 (£50.4M). Industrial research funding has grown more than fivefold, to £22.6M (£3.2M/year) compared to REF-2014. Research training income (UKRI CDTs) totals £22.1M. Additionally to Table 4, EPCC non-category-A staff have attracted £3.3M of research funding. Income to support impact is discussed in Section 3.4.

UKRI	£69.0M
EU H2020	£50.4M
Industry	£22.6M
Misc Overseas	£5.4M
Other	£20.0M
<i>Total</i>	£167.4M

Table 5: Research income by type over the REF-2021 period

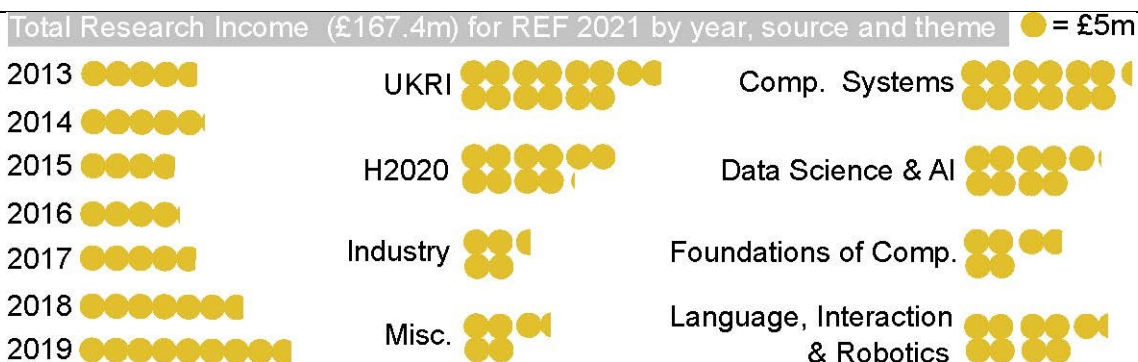


Figure 3: Research income (awards) over the REF-2021 period

All our impact case studies have arisen from externally funded research, and have underpinning research funded during the REF-2021 period. Research income is spread across the four themes:

- Computer Systems:** £64.2M (£9.2M/year), including EPSRC and RAEng fellowships to **O'Boyle**, **Leather** and Petoumenos. Major EPSRC grants, e.g. the Software Sustainability Institute (**Chue Hong**; £18.1M) and the ASiMoV prosperity partnership (**Parsons**; £8.1M), very large EU grants such as FORTISSIMO (**Parsons**; £8.0M), and industrially-funded PhD studentships (ARM, Cisco, Huawei, Microsoft, Synopsys, and Wolfson).
- Data Science and AI:** £26.3M (£3.8M/year) including fellowships from EPSRC (**Martin**) and the Royal Society (**Belle**; **Albrecht**). Funding from EPSRC, NERC, MRC, Wellcome, and EU H2020, including the NERC DAPHNE grant monitoring air pollution in Delhi using speckled computing (**Arvind**, £1.6M), a series of EPSRC grants on data-driven domestic energy monitoring (**Goddard**, £1.8M), and the EU Human Brain Project (**Armstrong**, £1.1M).
- Foundations of Computation:** £29.2M (£4.2M/year), including fellowships from EPSRC (**Heunen**, **Kashefi**, **Libkin**, **Sun**), ERC (**Cheney**, **Fan**), and the Royal Society (**Kammar**). Major industry funding from IOHK (**Kiayias**; £3.3M; Blockchain Technology Laboratory) and Huawei (**Fan**; £4.1M; database research). UKRI funding includes the VADA programme grant (**Libkin**; £1.5M), and the UK Quantum Technology Hub (**Kashefi**; £760k).
- Language, Interaction, and Robotics:** £47.7M (£6.8M/year), including fellowships from ERC (**Lapata**, **Steedman** (twice), **Titov**, Ferrari), UKRI (**Birch**, **Khadem**), and the Royal Society (**Komura**). Significant grants include three EPSRC Robotics and AI Hub Awards (**Vijayakumar**, **Mistry**, **Ramamoorthy**; £5M), and a series of large EPSRC grants in Speech Technology (**Renals**, **Bell**; £3.2M). Major EU H2020 funding includes 15 projects in machine translation and speech recognition (**Birch-Mayne**, **Germann**, **Haddow**, **Heafield**, **Renals**, **Sennrich**; £6.4M), as well as US government funding (£1.8M) from

DARPA (**Ramamoorthy**) and IARPA (**Bell, Heafield, Lapata, Renals**). Robotics research is also supported by the Edinburgh City Deal (£3.6M).

Benefits-in-kind include significant contributions from partners on UKRI grants – we secured cash and in-kind commitments of over £60M industry support for our five CDTs, enabling research projects and engagement, e.g. through CDT industry days and sponsored PhD projects.

3.2 Support for research

Our strategic goals include increasing interdisciplinary research and translational research and development. A key mechanism to enable research excellence is our research services infrastructure. Each research theme has at least two Portfolio Managers, who act as the first point of contact for staff seeking or receiving external research funding and are responsible for the management of all research grants and awards, including proposal preparation, grant negotiation, financial management and reporting, research staff contracts, and support for ethical and legal compliance. Our aim is to reduce administrative burdens on researchers and PIs, allowing them to focus on research.

For translational and commercial research, we have a team of 7 business development executives in the School of Informatics and Bayes, and a team of 15 R&D and project managers in EPCC, with the specific role to develop industrially-supported research and commercialisation. Our greater than 500% increase in industrially-funded research is an indicator of the success of this approach.

3.3 Research infrastructure and facilities

We are based primarily in three co-located award-winning buildings in the central George Square campus of the university: the Informatics Forum, the Appleton Tower, and the new Bayes Centre. In addition, we have a custom-built data centre, the Advanced Computing Facility (ACF), and those working in medical informatics are embedded in the Usher Institute within the Medical School.

The majority of UoA-11 staff are based in the Informatics Forum, a 12,000m² building opened in 2008, and the newly refurbished Appleton Tower, re-opened in 2017. The **Informatics Forum**, which has won 9 awards for architecture and for sustainability, houses 600 researchers including the majority of our PhD students. The **Appleton Tower** contains lecture theatres, social, study and commercialisation spaces, robotics labs, meeting rooms, and a pop-up cafe. The quality of its refurbishment was recognised through winning the **Design Through Innovation** category in

the 2018 RICS (Royal Institution of Chartered Surveyors) Awards. It also hosts a new cyber-security lab, workshop space, and a small maker space.

The £45M **Bayes Centre** (Figure 4), opened in 2018, is the university's innovation hub for Data Science and Artificial Intelligence, and is adjacent and physically connected to the Informatics Forum. The open-plan building won the **Large Project Award** at the 2019 Edinburgh Architectural Association Awards, and is designed for interdisciplinary, collaborative research. External organisations occupy about a third of the space in the building, enabling academic and industry researchers to “breathe the same air”. Bayes houses EPCC, the Edinburgh Centre for Robotics, the Huawei, RBS, and Rolls-Royce joint academic-industry centres, the DataLab Innovation Centre, and three of our CDTs (Robotics, NLP, and Data Science and AI).



Figure 4: The Bayes Centre.

The **ACF** is our custom-built data centre, with a focus on energy efficiency and automated monitoring and management of systems, situated on the outskirts of Edinburgh. It is the UK's largest and most capable data centre focused on supercomputing and research data systems, hosting national supercomputing facilities (Section 3.5).

All our buildings have been designed with accessibility in mind, including step-free access throughout, and disabled toilet facilities designed with input from wheelchair users based in the building.

Compute and Data Facilities: In addition to the national compute facilities described below, the university has an integrated network providing wired and wireless connectivity across its estate, with secure remote access to university computing, data and library resources. The School of Informatics extends this with its own security, network, storage and backup infrastructure, and a managed platform that provides a consistent Linux installation across more than 1,000 machines including over 300 servers. A staff of 26 dedicated computing/technical personnel support the computing facilities. Our mechanical and electrical workshops design, develop and maintain equipment for staff and student research projects.

The **ECDF Linux Compute Cluster** is the university's research compute cluster, and consists of some 10,000 Intel Xeon cores with up to 3 TB of memory available per compute node, and a

high performance, highly reliable filesystem designed to allow many thousands of running jobs to read and write to it simultaneously. It also includes 30 nodes for GPU computing, with a total of 152 GPU cards.

The School of **Informatics Compute Cluster** contains a further 175 compute servers, totalling about 4,000 cores, along with 160 GPU servers, with over 850 GPU cards. The cluster has about 2 petabytes of storage.

DataStore, the university's network file storage service, provides all researchers (staff and student) in the university a personal allocation of 500 GB, which can be pooled into shared group spaces.

Supercomputing is supported through EPCC. Our national level shared facilities are discussed in Section 3.5. EPCC also hosts local experimental systems which researchers can access, including the 4,096 Arm-core HPE Catalyst system, installed in 2018, and the Fujitsu NEXTGenIO system, which is the world's first large-scale HPC system to use Intel's Optane DCPMM non-volatile RAM technology.

Our suite of laboratories for research in **speech, language, and cognitive science** includes four sound studios (one hemi-anechoic) with equipment for synchronised ultrasound imaging and electromagnetic articulography, two **perception labs** with facilities for reaction-time measurements and audio presentation and recording, an **eye-tracking lab** with three high-speed eye-trackers, a **shape-capture lab** with a high-resolution static shape-capture system for 3D modelling, and an **instrumented meeting room** for multimodal recording of multiparty interaction. We have also established Design Informatics studio space, a Cyber-Security lab, and a Motion Capture lab. Our extensive equipment for robotics research is described below in the context of the national Robotarium.

3.4 Infrastructure for impact

Our organisational structure to enable and support the creation of impact from our research (Section 1.4) is mirrored by a financial and physical infrastructure.

- The Bayes Centre is central to the £237M Data-Driven Innovation programme which enables the university to collaborate with industry partners on public, private, and third sector projects, with the aim of fostering data-driven growth through translational research, large-scale data and compute infrastructure (the EIDF), and training.
- Informatics Ventures and the Informatics Commercialisation Team are based in the Bayes Centre. Funding for the team, and the programmes they organise (e.g., the Engage Invest Exploit (EIE) programme for early-stage data-driven innovation companies) is provided by Scottish Enterprise (£4.6M during the REF period). To date

EIE cohorts, comprising 237 young companies, have raised over £700 million in seed and later stage funding.

- EPCC transforms our research findings in supercomputing into commercial impact through (1) the HPC Accelerator programme which enables companies to take advantage of supercomputing, and (2) provision of consultancy and software development services to industry and commerce in areas including modelling and simulation, data science, on-demand computing, and HPC training. EPCC commercial income over the REF period is £3M/year, of which 2.2M/year comes from commercial research service contracts.
- The Bayes Centre AI and Blockchain Accelerator, sponsored by Telefonica and Cisco, was founded in 2018 (and launched by Prime Minister May). It hosts two annual cohorts of 10 start-ups, with the aim of enabling them to scale globally. The first cohort of start-ups secured £3.3M in private equity funding during the programme.
- Appleton Tower commercial incubation space is used by six University start-ups, including Actual Analytics, Cereproc, Quorate Technology, and SpeakUnique,
- Co-located industrial research teams in the Bayes Centre and Appleton Tower include Disney, Huawei, Intel, IOHK, Orbital Microsystems, and Royal Bank of Scotland.

We have established a 3-stage model to enable high-growth, data tech companies to be created, and supported from start-up seed funding through Series A and B investment rounds (Figure 5). At each stage, we provide a blend of data tech and entrepreneurship support.



Figure 5: Model to enable growth

3.5 Shared research infrastructure

We have a strong commitment to the national research infrastructure through supercomputing facilities, the Alan Turing Institute, and the national Robotarium.

We host and administer the majority of **national-level supercomputing facilities** for use by UK researchers on behalf of various UK research funding councils:

- ARCHER (4,920 compute nodes; 118,080 cores) and ARCHER2 (5,848 compute nodes; 748,544 cores) UK national supercomputing facilities provided by UKRI through a collaboration of EPSRC and NERC;
- STFC DiRAC UK national supercomputing facility (1,468 nodes; 35,235 cores);
- EPSRC Cirrus Tier 2 HPC facility (280 nodes; 10,080 cores, 152 GPUs), part of the Tier 2 HPC network of systems across the UK;
- EPSRC UK Research Data Facility (UK-RDF) UK national data facility (23PB disk).

In 2020, EPCC's ACF data centre, which hosts all of these systems, has been extended to include a new 6MW computer room and 30MW power supply in preparation for Exascale supercomputing in the next decade. The power supply is an £8M investment by the University, increasing our ability to host and operate supercomputers of international significance. The Edinburgh International Data Facility (EIDF), supported by the Edinburgh City Deal, will be located here: EIDF phase 1, completed in 2020, is a preliminary 150 node compute cluster with 20PB of storage.

Additionally, we are a partner in JADE, an EPSRC-supported Tier 2 facility, a national GPU service supporting world-leading research in machine learning. We have also made extensive use of the CSD3 national GPU facility, via UKRI funding.

The new national **Robotarium** (co-director *Vijayakumar*) expands on the previous facilities of the Edinburgh Centre for Robotics and operates as a national UK facility for industry and academia. The Robotarium has created a unique integrated capability to explore collaborative interaction between remote teams of humans, robots and their environments at all levels. Equipment in the Bayes Centre facilitating this aim includes: the NASA Valkyrie and the PAL Talos humanoid robot (Figure 6); the PR2 and Baxter two-armed robots; the Husky unmanned vehicle; the ANYmal autonomous four-legged robot; and Living Labs which mimic real domestic, commercial and medical situations in which robots could be deployed. In addition, there is a wide range of enabling equipment including rapid prototyping and micro-assembly equipment, sensor systems and on-board computing.

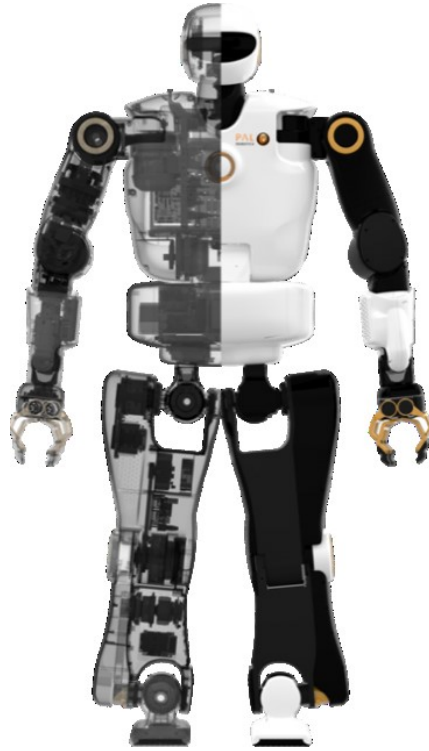


Figure 6: The PAL Talos humanoid robot

4. Collaboration and contribution to the research base, economy and society

4.1 Research collaborations, networks, and partnerships

We have strategic collaborations at international, national, and regional levels. Internationally, most of our collaborations arise from research projects funded by the EU, by industry, by charitable foundations, and by national agencies for example from the USA (IARPA, DARPA), Japan (JST), and Switzerland (SNF). We are also part of the **UNA-Europa alliance** of eight leading European research universities which has a research focus on data science, sustainability, and cultural heritage. We are an ELLIS Unit in **ELLIS**, the European Laboratory for Learning and Intelligent Systems, members of the **Multilingual Europe** initiative (and a funded partner supporting the European Language Grid), founding members of the European **HiPEAC** network for computer systems research, lead the UK's involvement in the **PRACE** research infrastructure and the **HPC-Europa** network for supercomputing, and a funded partner of the **Human Brain Project**.

We are a founder university partner of the **Alan Turing Institute (ATI)**, the UK's national institute for Data Science and AI (lead: **Williams** (2016-2018), **Rovatsos** (2018-date)), within which the university leads the Artificial Intelligence (co-director: **Vijayakumar**) and Finance and Economics programmes. 20 of our category-A staff are Turing Fellows, with a further 28 Turing Fellows returned in the university's submissions to other UoAs. The Bayes Centre houses the only physical space of the ATI outside London.

We take a leading role in **SICSA**, the Scottish Informatics and Computer Science Alliance, current director **S Anderson**. In addition to providing a networking infrastructure across several hundred computer science and informatics researchers in Scotland, SICSA supports an annual PhD student conference, research communities across seven specific research themes, a programme of distinguished visiting fellows (DVF's), and a knowledge exchange programme, centred on the annual DemoFest, which brings together university researchers with private and public organisations. SICSA was also instrumental in establishing the Data Lab, Scotland's innovation centre for data science and AI. During the REF period, the SICSA DVF programme has supported research visits from 11 leading international researchers hosted at Edinburgh, such as Jeremy Siek (Indiana University), Jimmy Lin (Waterloo University), and Esra Erdem (Sabanci University).

Across the REF period, we have hosted 575 academic visitors (over 70 per year), and reserve office-space and facilities for such visitors. Recent visitors funded by external fellowships include David Dunson (Carnegie Trust Centenary Professor), and Daniel Sorin (Leverhulme Visiting Professorship). Our visitor programme has led to significant research collaborations, as evidenced by joint publications such as a paper by Sorin and **Nagarajan** was an IEEE Micro top pick honourable mention, and a series of papers from **Wadler** and visitor Jeremy Siek.

Our programme of **Distinguished Lectures**, including the prestigious annual Milner and Oberlander Lectures, attracts world-leading researchers, with 4-6 lectures being held annually. Examples of distinguished lecturers in the REF period include Turing Award winners Yann Le Cun and Silvio Micali, as well as Tom Dietterich, Barbara Grosz, Dexter Kozen, Kathleen McKeown, Mari Ostendorf, Ben Shneiderman, J Strother Moore, and Moshe Vardi. During 2020 we organised a lecture series in Ethical AI, Data Science and Algorithms with speakers including Shannon Vallor and Moritz Hardt. Our distinguished lectures have continued online since March 2020, including the 2020 Milner Lecture.

Nearly all our staff have **international research collaborations** – 96% (158/164) of our category-A staff published papers with international co-authors, and our co-authors come from 49 of the top 50 universities in computer science (as ranked by THE 2020).

During the REF period we have participated in 102 collaborative EU projects, having coordinated projects including the FORTISSIMO and FORTISSIMO-2 supercomputing projects (£11.7M; **Parsons**; total projects value €33.4M), the SUMMA media monitoring project (£1.3M; **Renals, Birch-Mayne, Cohen**; €8.0M), the PANORAMIX privacy-preserving communications project (£343k; **Kiayias**; €4.5M), the THING robotics project (£659k; **Mistry**; €4.1M), the ESSENCE Marie Curie Training Network (£950k; **Rovatsos**; €4.0M), the ADEPT HPC project (£862k; **Weiland**; €3.6M), and the MINIMAL robotics project (£824k; **B Webb**; €3.0M). We have had collaborations with BBN, Columbia University, the University of Maryland, and Yale University supported by IARPA (£1.4M; **Bell, Heafield, Lapata, Renals**), and with Arizona State University, the University of Michigan, and Xerox PARC supported by DARPA (£1.2M; **Cheney, Grot, O'Boyle, Ramamoorthy**).

Significant EPSRC-funded collaborations include the Software Sustainability Institute (**Chue Hong, Parsons**; with Manchester, Oxford, and Southampton), the ASiMOV EPSRC Prosperity Partnership (**Parsons, Weiland**; with Bristol, Oxford, and Warwick, in partnership with Rolls-Royce), and the programme grants ABCD (**Wadler**; with Glasgow and Imperial), COG-MHEAR (**Bell, Renals**; with Edinburgh Napier, Glasgow, Heriot-Watt, Manchester, Nottingham), NST (**Renals**; with Cambridge and Sheffield), SOCIAM (**Buneman, Robertson**; with Oxford and Southampton), and VADA (**Libkin, Fan, Buneman, Pieris**; with Oxford and Manchester).

Diverse networks are a key outcome of our interdisciplinary research. The Creative Informatics Network (**Rovatsos**) combines research in design, data science, and social science with an industry network spanning banks, galleries, broadcasters, and publishers. **J Robertson** (with Glasgow, Stirling, and Abertay) has shaped computer science education in schools across Scotland through the construction of the teachcs.scot network, resulting in changes to educational and pedagogical practices and methods across 2,514 primary schools in Scotland. **Renals** and **Ribeiro** (with speech therapy at Strathclyde) have developed speech technology for

speech and language therapy, resulting in partnerships with four NHS trusts (Lothian, Glasgow, Grampian, UCL Hospitals). In collaboration with the SciStarter platform, **Gal** has used AI algorithms to enrich a citizen science network.

EPCC is the hub of several **interdisciplinary networks** based on its shared facilities (Sections 3.3 and 3.4), including the UK Research Data Facility, a large-scale data store supported by EPSRC and NERC and primarily used by researchers in material science, climate science, physics, engineering, and biosciences, and the ARCHER and ARCHER2 supercomputing facilities used by academic and commercial users (582 average monthly users, 354 industrial users in REF period). EPCC has also developed a network of active collaborations in data science, including a multidisciplinary study on the human impact of industrial revolution (with the Turing, British Library, and four universities – Cambridge, UEA, Exeter, QMUL), medical imaging projects (with Dundee, Abertay, and NHS Scotland), and supplying a safe data haven to Health Data Research UK (MRC), linking 22 universities and research institutes.

We have a strategic focus on addressing **global challenges**. Through the NERC-funded project DAPHNE (£1.6M) **Arvind** has established a network (with IIT Bombay, the University of Delhi, the Centre for Chronic Disease Control, and Indian Institute of Tropical Meteorology) addressing air pollution in Delhi using novel personal devices that analyse air quality and monitor breathing rate and effort. In the EPSRC project UNMUTE (£970k), **Bell** partnered with the NGO Translators Without Borders to develop spoken interaction systems in low-resource, but widely spoken, languages in India and South Africa.

Our collaborations lead to considerable **impact**. Five of our nine impact case studies grew out of collaborative research projects: the audio animation startup Speech Graphics is based in part on the EU Network of Excellence SSPNET (**Renals, Shimodaira**); the speech recognition startup Quorate has exploited research done in the EU projects AMI and AMIDA, and the EPSRC programme grant NST (**Bell, Renals**); the open-source software for neural machine translation arises from a linked set of EU projects (**Birch-Mayne, Haddow, Heafield, Sennrich**); the NextGenIO persistent memory layer is a direct output of an EU H2020 project (**A Jackson, Parsons, Weiland**); and the Ouroboros cryptocurrency protocol grew out of the EU PANORAMIX project and the ERC project CODAMDA (**Kiayias**).

4.2 Contributions to the economy and society

The impact of our research has increased significantly during the REF period. Of the impact case studies submitted to REF-2014, two have been extended to form the basis of REF-2021 impact case studies (Speech Graphics and Embedded Processors). Other impact case studies from REF-2014 have demonstrated lasting and sustainable impact:

- XML standards continue to have tremendous influence based on technologies in the previous impact case study: Office Open XML (ISO/IEC 29500-4:2016) uses XML

Schema and has defined the formats for Microsoft Office since 2013; EPub (ISO/IEC TS 30135), the e-book format of choice, uses XML Schema to define its structure, with most publishers using XSLT in their workflow.

- the NLTK toolkit continues to develop with regular new versions, and is now used in NLP courses at over 80 universities in 25 countries;
- the Tegola rural broadband project has expanded into the HUBS not-for-profit membership organisation of community network operators, combining 11 community networks;
- MilePost GCC continues to be the standard in machine learning compilation and is the basis for the MAGEEC compiler from EMBECOSM;
- commercial and clinical applications of speech synthesis, have continued to develop through their use by motor neurone disease patients in communication devices, and the recent commercialisation of speech synthesis for assistive technologies by SpeakUnique;
- the analysis and modelling of laboratory animal behaviour commercialised by a startup (Actual Analytics) and in use at companies including AstraZeneca, GSK and Bayer.

Each of our research themes have produced impact that goes beyond the 2014 and 2021 impact case studies.

Further examples of impact arising from our research in **Computer Systems** include an approach to optimise server usage deployed by Microsoft across its data centres; the characterisation of security vulnerabilities in Fitbit devices; work with HPE, ARM, Marvell and SUSE to host and operate a 4,096-core HPC system to explore the use of the ThunderX2 Processor for ARM-based HPC. **Chue Hong** co-authored the OECD Science, Technology and Industry Policy Paper "Building digital workforce capacity and skills for data-intensive science", which informs the work and policy of members of the OECD Global Science Forum.

Our research in **Data Science and AI** has made considerable impact through the RESpeck and AIRspeck devices (**Arvind**) including public health studies into air pollution in India. Our work on machine learning, predictive modelling, and autonomous agents provides the foundation for the design of driverless cars being developed by Five AI. Our data analytics software developed for computational neuroscience (Bednar) underpins the Anaconda data science platform. Former PhD and postdoctoral researchers have carried forward our vision for societal and industrial impact: Tom Griffiths was co-founder of FanDuel (valued at over US\$1 billion before acquisition), Jono Millin is co-founder of DroneDeploy, and Daniel Winterstein is co-founder of Good-Loop.

In **Foundations of Computation**, we have developed radical new algorithms for querying big data with bounded resources deployed across Huawei's big data platform, and for parallelised graph computations deployed by Alibaba (**Fan**). **Libkin** and **Guagliardo** have established a joint research team with the world leader in graph databases, Neo4j Inc, and sit on the ISO

committee that produces the SQL Standard. Combined with work on dependent types (**Plotkin**), work on the effect handler abstraction (REF-2021 impact case study) has also led to Microsoft's Everest project to design and implement a high performance, standards-compliant, formally verified, drop-in replacement of the full HTTPS ecosystem. Further impact includes Imandra, founded by Grant Passmore (former PhD student and postdoctoral researcher with **P Jackson**), a cloud-scale automated reasoning system used in the financial technology industry.

In **Language, Interaction, and Robotics** our research has resulted in the deployment by Hitachi of robotic intelligent warehousing (**Vijayakumar**), a teaching model to treat computer science as a scientific discipline, underpinning the new Scottish computing curriculum, now in use at over 2,500 schools across Scotland (**J Robertson**), and natural language processing and machine translation commercialised for personal finance by Aveni (**Birch-Mayne, Haddow**).

We collaborated with the National Museum of Scotland to produce the highly successful "Robots" exhibition from January–May 2019, totalling 61,002 visitors, which included a section detailing our current research, featuring the Valkyrie project with NASA and work on prosthetic limbs.

Other public engagement activities include our work on human-data interaction (Oberlander†, **Hill**) resulting in the Harmonium Project which opened the 2015 Edinburgh International Festival with a live audience of 20,000 people in Edinburgh city centre, and the Zizi artwork exploring machine learning and sexual identity, exhibited in our Informatics-based exhibition space, InSpace during the Edinburgh Fringe Festival 2019.

In June 2017, we partnered with Scottish Enterprise (SE) to deliver the £5.7M Bayes Innovation Programme (BIP), based in the Bayes Centre. BIP has supported company engagement and entrepreneurship to generate corporate collaborations, create and scale high-growth ventures, attract new investment and deliver economic benefit through:

- a dedicated business development team focussed on offering multi-disciplinary R&D opportunities with industrial partners;
- establishing an accelerator in AI and Blockchain with industry partner Telefonica through their subsidiary WAYRA and sponsorship by CISCO;
- the Informatics Ventures team attracting record sponsorship and attendance for the Engage Invest Exploit (EIE) Investment Programme;

The economic deliverables from BIP have greatly exceeded those planned:

OBJECTIVE	Target by Dec 2020	Actual at Dec 2020
Attracting 15 corporate R&D teams in or near Bayes	15	51
Securing a minimum of 90 new jobs by attracting these corporates	90	369
Establishing 6 additional high-growth ventures by July 2019	6	15
Attracting £45M of investment in high-growth companies	£45M	£92.8M
Generating £17M to support research, innovation and entrepreneurship	£17M	£38.1M

Table 6: BIP objectives

4.3 Contributions to the research base

Our staff have major research leadership roles. **Parsons** was recently been seconded (1 day/week) as EPSRC's Director of Research Computing – a role responsible for developing the UK's Digital Research Infrastructure for science and engineering. **D Robertson** was a member of the EPSRC Strategic Advisory Team (SAT) for ICT until 2017, and **Chue Hong** and **Weiland** are members of the EPSRC SAT for e-Infrastructure. In addition, a majority of our category-A staff serve as reviewers and panel members for main funding councils and research charities in the UK, Europe and internationally (UKRI, RAEng, ERC, RS, NSF, etc.), **Martin** was a founding member of the Royal Society Diversity Committee (2015-2018), and we contributed to Royal Society working groups: **Buneman** (Dynamics of Data Science Skills), **Bundy** (Explainable AI), and **Lapata** (Machine Learning). **Bundy** was chair of the Royal Academy of Engineering panel on Computing and Communications (2014-2017). **Rovatsos** and **Vijayakumar** sit on the Research and Innovation Advisory Committee of the ATI; **Vijayakumar** is the programme co-director for AI at ATI. **Oyarzún** was awarded a Fellowship from the World Economic Forum to join their global council on emerging technologies.

We have leading roles in the UK Computing Research Committee (UKCRC), the expert panel of the IET and BCS for computing research in the UK. **Hillston** is the current chair (2018–date) and **D Robertson** was chair from 2014-2016. 16 of our category-A staff are UKCRC members. We are an institutional member of Informatics Europe: **Hillston** was an Executive Board member (2014–2017) and was Chair of Women in Informatics Research and Education (2016–2017).

Professional service: Within ACL, **Goldwater** is European chair (2019-date), **Steedman** and **Lapata** were SIGDAT presidents in 2016 and 2018. **Williams** is a member of the NeurIPS Foundation advisory board and **Renals** has served as an IEEE SPS Technical Committee member. In the REF period 27 of our staff have served as senior editors of 52 international journals (e.g. Logical Methods in Computer Science, Journal of Open Research Software and IEEE Open Journal on Signal Processing), and 49 are editorial board members of 99 international journals. 52 of our staff have acted as chair for 119 conferences (e.g. AISC, ACL, CGO, IJCAI, LICS, ParCo, PKC, PLDI) and workshops, and our staff have served on the programme or organising committee for over 600 international conferences and workshops.

Fellowships of learned and professional societies awarded during the REF period: Royal Society (**Fan, Bishop**); Royal Society of Edinburgh (**Lapata, Martin, Oberlander†, Sannella**); RSE Young Academy Scotland (**Belle, McNeill, Ramamoorthy, Simpson**); Royal Academy of Engineering (**Martin**); ACM (**Bundy**); IEEE (**Renals**); ACL (**Lapata**); ISCA (**Renals**); Academia Europeae (**Buneman, Fan, Hillston, Lapata, Libkin**). As stated in Section 2, 23 of our researchers were awarded competitive fellowships funded by EPSRC, ERC, the Royal Society, RAEng, and UKRI during the REF period.

Awards and prizes include: the BCS Roger Needham Award (**Goldwater** 2016), Ada Lovelace Medal (**Plotkin**, 2018), and Karen Spärck Jones Lecture (**Lapata**, 2019); the ACL Lifetime Achievement Award (**Steedman**, 2018; Webber 2020); Royal Society Wolfson Research Merit Awards (**Fan, Lapata, Libkin**); the ACM SIGMOD Test-of-Time Award (**Bunemen, Cheney**, 2016); the Alberto O. Mendelzon Test-of-Time Award for ACM PODS (**Fan**, 2015); the CADE Influential Paper Award (**Bundy**, 2017); the CGO Test of Time Award (**O'Boyle** 2017; **Leather, O'Boyle**, 2019); the Eurographics Young Researcher of the Year (**Bach**, 2019); and BCS Distinguished Dissertation Awards in 2014 (Reutter, supervised by **Libkin**) and 2018 (Allamanis, supervised by **Sutton**). Our researchers have won over 60 best paper awards at international conferences including ACL, ACM SIGMOD, ECCV, ICALP, ICDT, ICML, ICS, Interspeech, KR, and VLDB, and have given over 300 invited keynote talks (e.g. ACL, FLOC, PODS, IEEE HSCMA, ISPASS, POPL, WCAI).