



UNIVERSITY OF LEEDS

LEEDS *Institute for Data Analytics*

ANNUAL REPORT ACADEMIC SESSION 2016 / 17



lida.leeds.ac.uk

Contents

- 2** Welcome

- 4** LIDA Director's Report

- 6** LIDA Senior Management Team

- 8** LIDA Funding

- 9** LIDA Governance

- 10** LIDA Research & Innovation

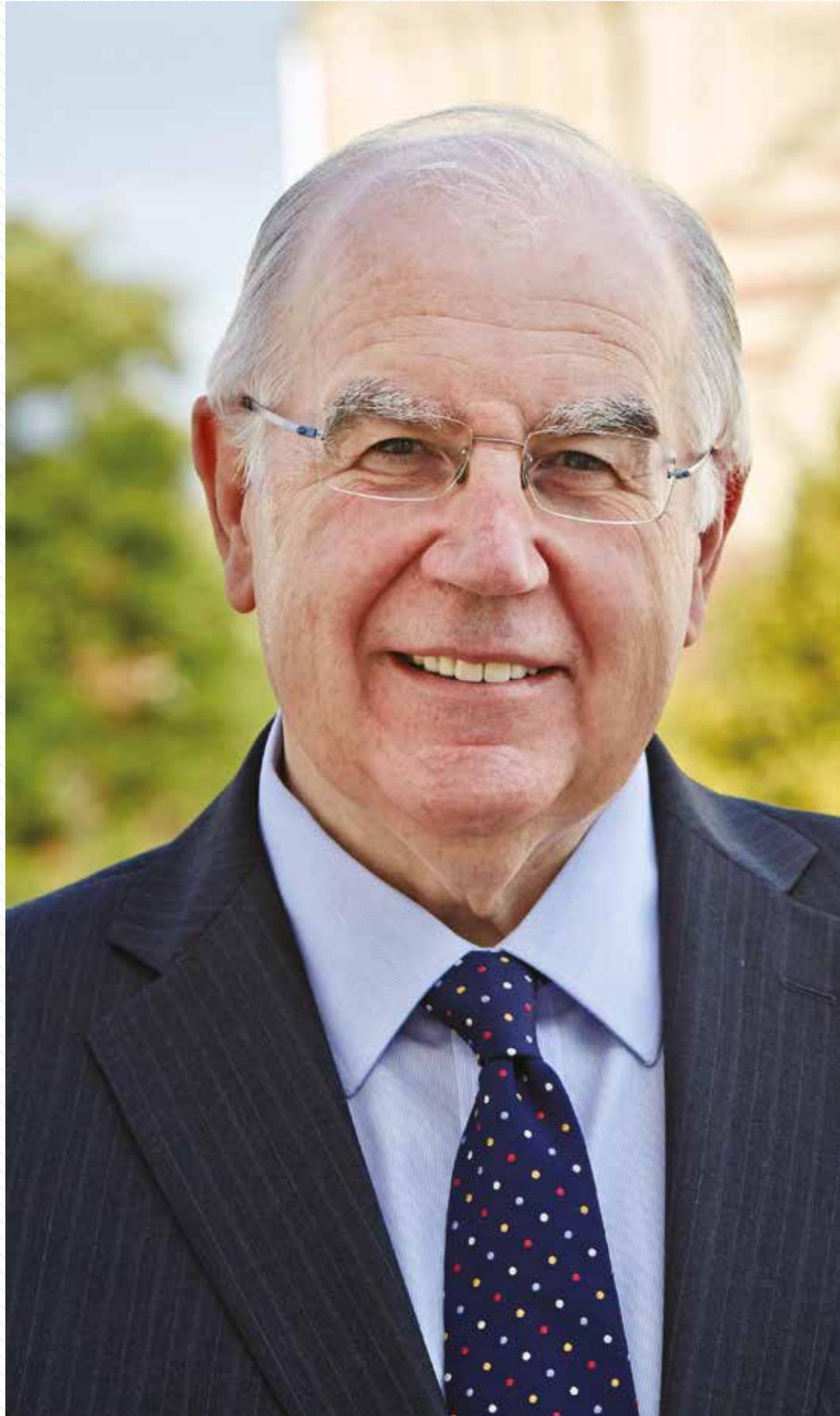
- 12** MRC Medical Bioinformatics Centre

- 14** ESRC Consumer Data Research Centre

- 16** LIDA Education and Training

- 17** Integrated Research Campus

- 18** LIDA Partnerships



Welcome

At a time of great change in the world, there is a national imperative to create an innovative data environment in the UK that will sustain the excellence of the research base, increase productivity and transform the way in which public and private sector organisations do business and provide services.

The Leeds Institute for Data Analytics (LIDA) is at the heart of this, developing data analytics as a transformative technology and playing a crucial role at the intersection between research, education and effective partnership working with a wide range of industry partners.

By bringing together applied research groups with data scientists from mathematics, statistics and computer science, and with the ESRC Consumer Data Research Centre and the MRC Medical Bioinformatics Centre at its core, LIDA is opening up new ways of understanding health and human behaviour and casting light on a wide range of social and environmental problems.

LIDA is housed in a purpose built facility which includes data safe rooms, advanced computational infrastructure and a fully equipped training suite. Crucially, it benefits from the University's Integrated Research Campus (IRC), a unique initiative which provides a secure data handling platform that transcends disciplines, institutions and economic sectors, enabling new and exciting models for collaborative working.

LIDA is already a success story. This report shows that: the current value of research projects is £45m, including grants from all of the research councils and major charities; we have a sector leading approach to training and capability building which spans undergraduate students to a strong international group of early career researchers and features a highly prized data scientist intern programme; and that our engagement with a wide range of external partners, including national organisations based in Leeds (eg NHS Digital), major international corporates (eg Asda-Walmart and Aviva) and key statutory organisations (eg ONS and HMRC), is highly developed.

In a very short time, LIDA has become an important national asset as a leading research and development institute with the capacity to drive many aspects of the UK Industrial Strategy and to enhance our work in low and middle income countries through the Global Challenges Research Fund. I congratulate the leadership team at LIDA on all that they have achieved so far and will ensure that the University continues to support their medium and long term ambitions.

Sir Alan Langlands

FRSE HonFMedSci

Vice-Chancellor, University of Leeds

LEEDS *Institute for
Data Analytics*

LIDA DIRECTOR'S REPORT

Business, Government and Society will be transformed in the coming years by the ways in which data are captured, processed and analysed.

In the 2016 best seller, *Homo Deus: A Brief History of Tomorrow*, Yuval Noah Harari cites recent research that suggests 95% of chefs could be replaced by algorithms in the year 2033.

With a degree of variation and some notable exceptions, other jobs and professions are heading in the same direction. Handled correctly, such changes could generate hugely positive social and economic benefits as well as impacts on employment and business competitiveness.

If such predictions are only halfway correct, then data analytics promises to be the transformative technology of the 2020s, with universities playing a crucial role at the intersection between research, education and knowledge transfer. In creating the Leeds Institute for Data Analytics, the University of Leeds has taken a bold step towards leadership in this important domain.

Thanks to the support of senior colleagues in the University's Executive Group, LIDA can boast the highest quality infrastructures and technical support, a physical home and, from this year, a shared contribution model with every Faculty combining resources and benefiting from our successes.

This report showcases how well LIDA provides the conditions for a potent mixture of multidisciplinary research and relevance to end-users. Our work is funded by the Research Councils, Innovate UK and medical research charities, with our PhD students benefitting from partnerships with public and private sector organisations, both local and global.

The National Research Centres in Consumer Research (CDRC) and Medical Bioinformatics (MBC) – which create a distinctive collaboration at the heart of LIDA – have both seen continued progress this year. CDRC won the competition to establish the UK's first Centre for Doctoral Training in Data Analytics and Society and our four year Integrated Programme aims to set new standards for PhD research in data science, not least through its strong engagement with real world partners. Our medical colleagues won an important research award from Cancer Research UK for the creation of a unique health care intelligence hub focusing on bowel cancer data.

September 2016 saw the start of a highly innovative intern scheme which is perhaps unique in UK Data Science. The competition was fierce, with more than 120 applications for just ten places. Supervised by LIDA academic partners, our interns undertook projects ranging from waste collection to mass mortality of antelope in Central Asia and have now moved on to new jobs or, for some, PhD studentships within LIDA. We will welcome our next intake in September 2017.

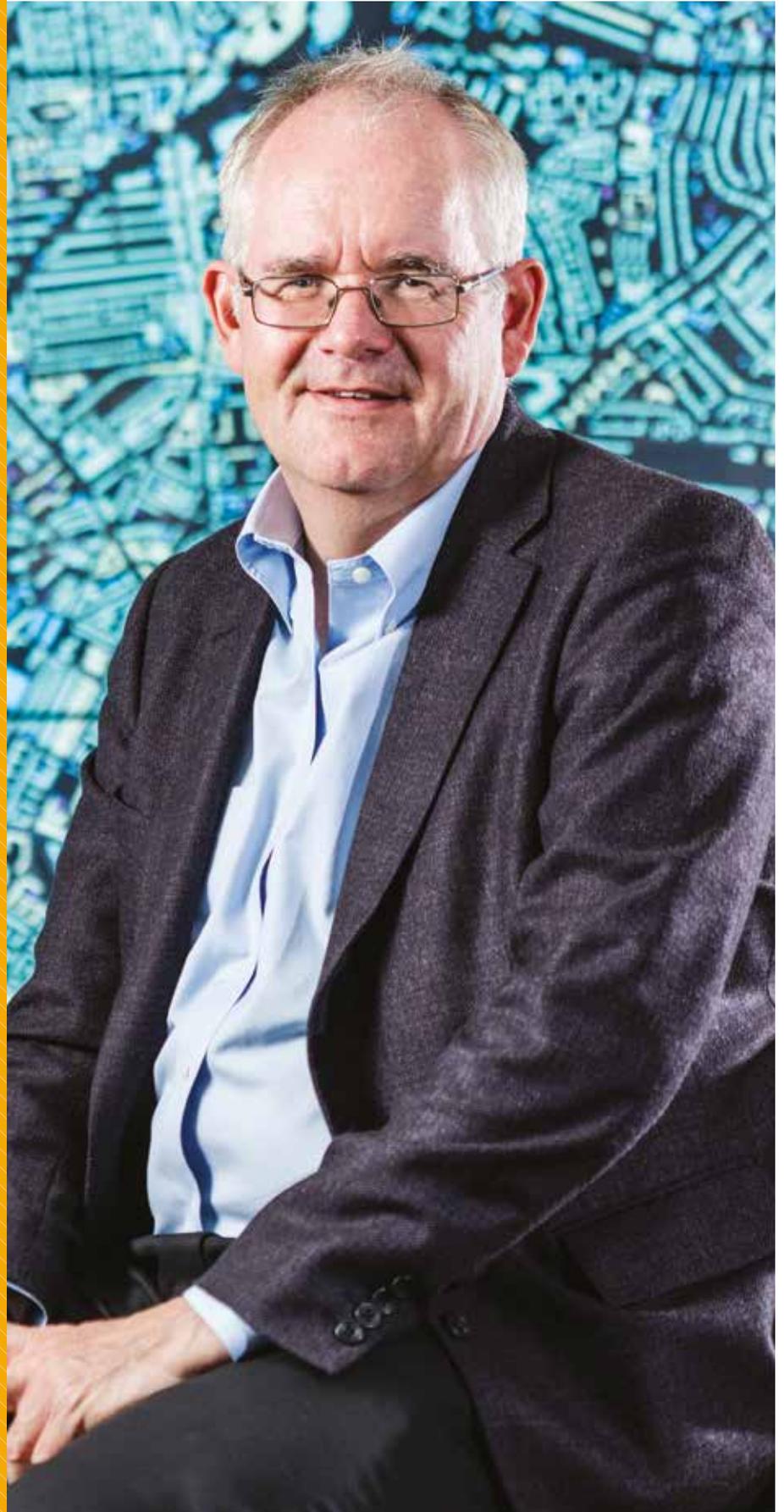
This spring, with the help of colleagues in the schools of Business, Mathematics, Computing, Medicine, and Geography and in Research Innovation and IT, LIDA's Senior Management Team undertook a strategic review. While the outcomes remain subject to validation with partners and colleagues, I feel confident that we will retain a steady focus on core values of multi-disciplinarity, real world engagement, excellent infrastructure and a genuinely collegial and collaborative approach.



“Whether you are reading this report as an academic, an external partner, a potential student or in some other capacity, then if you share in our values and combine excitement about, with experience in, the power of data, I look forward to working with you!”

In the last 12 months, we have significantly expanded our data services team to develop and deliver the Integrated Research Campus (IRC) which is integral to our work. This team achieved ISO accreditation for their secure operating models in May 2017, which were assessed independently through research partners including NHS Digital and audited by PwC.

We now have an analytics platform which is the envy of many academic collaborators, competitors and business partners, and look forward to developing this further.



LIDA SENIOR MANAGEMENT TEAM

The LIDA Senior Management Team supports the LIDA Director in executive decision-making and operational matters and is responsible for:

- determining and delivering LIDA's Business Plan
- managing LIDA's financial performance
- ensuring engagement with external partners and agencies
- overseeing operational management of LIDA staff, finance, information, IT, estates and health and safety
- monitoring the effective implementation of Information Governance processes and procedures.



Professor Mark Birkin, BA, PhD
LIDA Director, ESRC Consumer Data Research Centre Director and Professor of Spatial Analysis & Policy, School of Geography
m.h.birkin@leeds.ac.uk



Professor Sir Alex Markham,
MBBS, BSc, PhD, DSc, FRCPath, FRCP, FMedSci
LIDA Deputy Director, MRC Medical Bioinformatics Centre Director & Professor of Medicine, School of Medicine
a.f.markham@leeds.ac.uk



Adrian Iredale, BA, MBA, MA, ALCM
LIDA General Manager
j.a.iredale@leeds.ac.uk

Manages the strategic development and direction of LIDA covering multi-disciplinary and world-leading research, student education, infrastructure, finance, HR, external relationships and stakeholder engagement.



Philip Waywell, MChem, MSc, PhD
LIDA Research & Innovation Manager
p.d.waywell@leeds.ac.uk

Responsible for the continued growth of LIDA's Research & Innovation portfolios and contributes to external relationship and stakeholder engagement activities.



Andy Pellow, BA, MSc
Acting Head of Business Relationship Manager, University IT Service
a.j.h.pellow@leeds.ac.uk

Responsible for the IT Service's strategic support for LIDA.



Dr Geoff Hall, BMedSci, MBChB, PhD, FRCP

Senior Lecturer, Medical Oncology & Associate Medical Director – Informatics, Leeds Teaching Hospitals NHS Trust
g.hall@leeds.ac.uk



Thomas Fleming, MA (Cantab), MInstP, MCITP

IRC Data Services Manager, IT Services
t.j.fleming@leeds.ac.uk

Manages and maintains the Integrated Research Campus secure virtual research environment and assists researchers all at stages of the research life cycle with the use of sensitive data for research purposes.



Gary Cartwright, ACMA, CGMA

LIDA Finance Manager
g.cartwright@leeds.ac.uk

Responsible for monitoring and reporting on the financial performance and sustainability of LIDA to the Senior Management Team.



Amy O'Neill, MA, MA

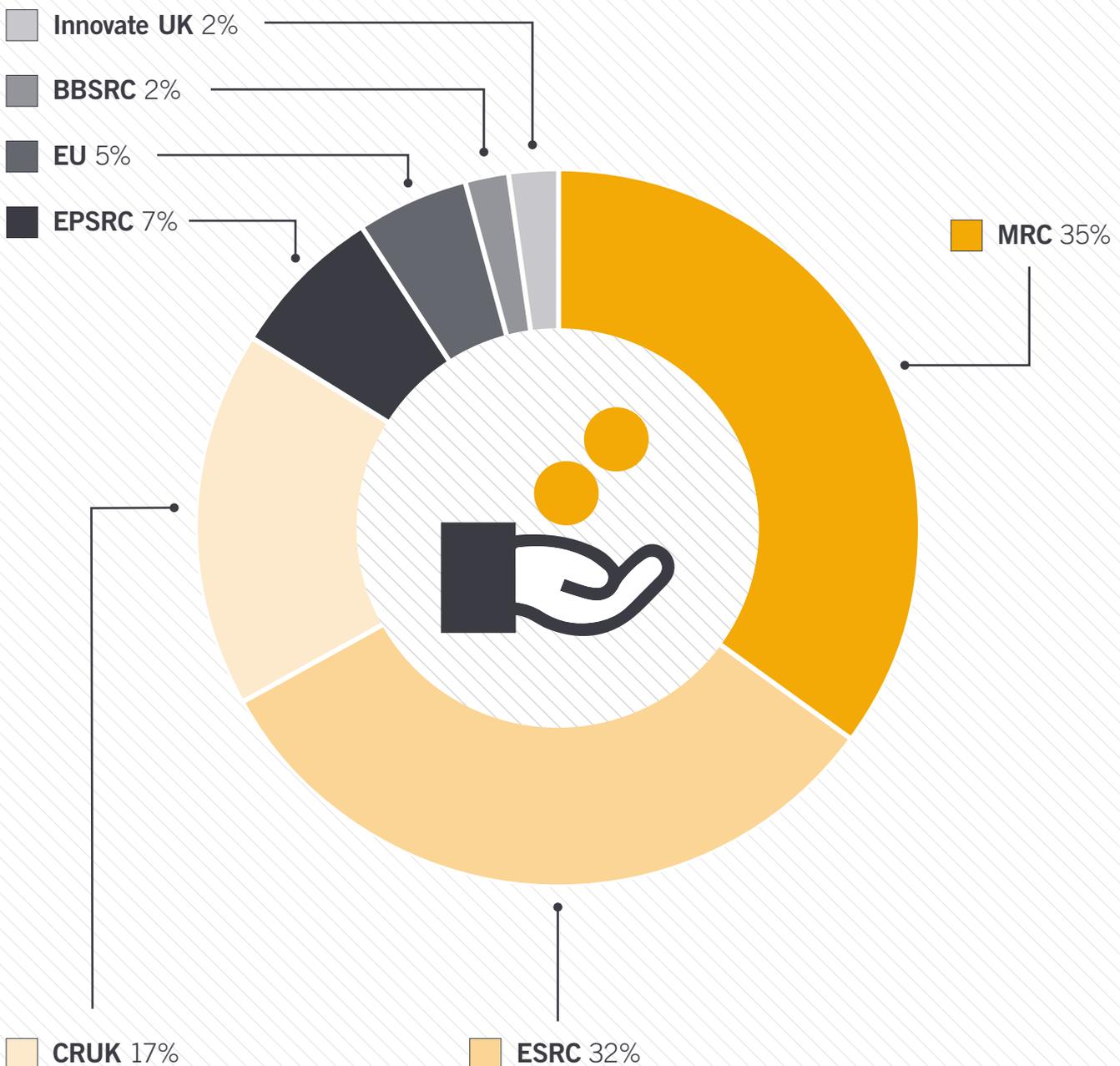
Manager, ESRC Consumer Data Research Centre
a.oneill@leeds.ac.uk

Responsible for the strategic development and the day-to-day management of the ESRC-funded Consumer Data Research Centre's programmes and services.

LIDA FUNDING

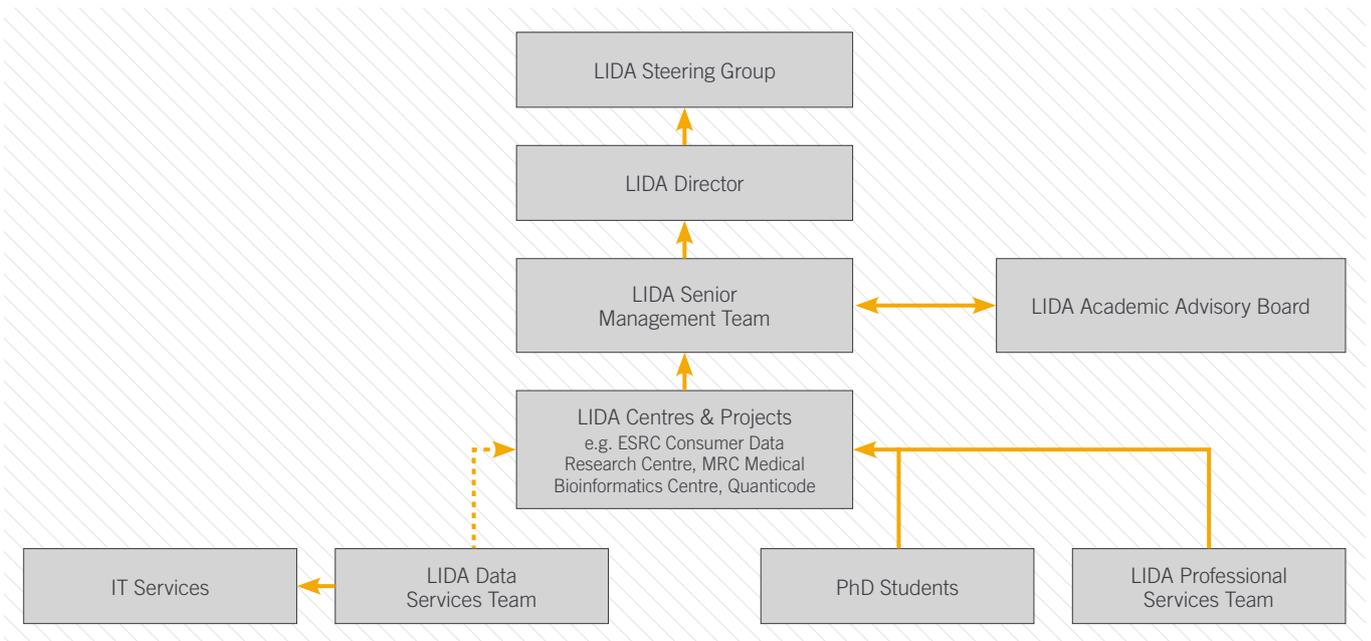
This distribution of funding is based on the first ten projects in LIDA, including the two foundation centres, namely the **MRC**'s Medical Bioinformatics Centre and the **ESRC**'s Consumer Data Research Centre together with the following eight additional first wave projects that are co-located in LIDA:

CRUK UK Colorectal Cancer Intelligence Hub, **ESRC** Centre for Doctoral Training in Data Analytics and Society, **ESRC** Simulating Urban Flows, **EPSRC** Quanticode, **EPSRC** Multi-Scale Infrastructure Transitions, **Innovate UK** Citizens at the City's Heart, **EU** Empower and **BBSRC** PigSustain



LIDA GOVERNANCE

LIDA's Senior Management Team (see pages 6-7), Academic Advisory Board and Steering Group have joint responsibility, on behalf of the University of Leeds, for governance of the Institute.



The **LIDA Academic Advisory Board** is responsible for determining LIDA's research & innovation, student education and training priorities and maintains oversight of:

- LIDA's key performance indicators in respect of the quality and quantity of research and student education outputs
- potential sources of funding and opportunities for commissioned research
- effective networks across University faculties regarding issues of mutual academic benefit.

Members:

- Professor Mark Birkin, LIDA Director (Chair), Faculty of Environment
- Professor Alexis Comber, School of Geography, Faculty of Environment
- Professor Jocelyn Evans, School of Politics and International Study, Faculty of Education, Social Sciences and Law
- Professor Christopher Forde, Leeds University Business School

- Barry Haynes, IT Enterprise Architecture, IT Services
- Professor Jeanine Houwing-Duistermaat, School of Mathematics, Faculty of Mathematics and Physical Sciences
- Adrian Iredale, LIDA General Manager
- Professor Sir Alex Markham, Director, MRC Medical Bioinformatics Centre, Faculty of Medicine and Health
- Professor Christopher Megone, School of Philosophy and History of Science, Faculty of Arts, Humanities and Cultures
- Professor Roy Ruddle, School of Computing, Faculty of Engineering
- Dr Philip Waywell, LIDA Research & Innovation Manager
- Professor David Westhead, School of Molecular and Cellular Biology, Faculty of Biological Sciences
- Professor Stephen Westland, School of Design, Faculty of Arts, Humanities and Cultures
- Dr Ceri Williams, Director of Research Innovation & Development

The **LIDA Steering Group** determines LIDA academic strategy in line with University strategy and objectives.

Members:

- Professor Andrew Dougill, Executive Dean, Faculty of Environment (Co-chair)
- Professor Paul Stewart, Executive Dean, Faculty of Medicine and Health (Co-chair)
- Professor Mark Birkin, LIDA Director
- Professor Andrew Gouldson, Dean, Interdisciplinary Research, Faculty of Environment
- Professor John Ladbury, Executive Dean, Faculty of Biological Sciences
- Jane Madeley, Chief Financial Officer, University of Leeds
- Professor Sir Alex Markham, Director, MRC Medical Bioinformatics Centre
- Professor Lisa Roberts, Deputy Vice Chancellor (Research & Innovation)

LIDA RESEARCH & INNOVATION

LIDA brings together over 150 researchers and data scientists across the University to develop new approaches to understanding health and human behaviour and tackle a diverse range of social and environmental problems.

LIDA has a distinctive emphasis on applications of data science across a wide variety of academic domains in partnership with user organisations from business, government and the third sector. The Institute provides a unique collaborative environment that supports the formation of new, interdisciplinary research teams with the right combination of skills needed to address real world challenges.

The University has made a significant investment in academic and technical staff working within the LIDA portfolio in recent years, including: 6 Professors; 36 University Academic Fellows; and an eight-strong Data Services Team.

“Sixteen research centres, programmes and projects are now based at the Institute and our infrastructure also supports £45 million research across the university.”

Researchers:

150



University Academic Fellows:

36



Data Services team:

8



LIDA research portfolio

2017

£45
MILLION

Key strengths

Health Informatics



Consumer Data Research



Transport and Mobility



Crime Analytics



Data visualisation



Statistical and mathematical methods



National Centres of Excellence hosted at LIDA

- MRC Medical Bioinformatics Centre (see page 12)
- ESRC Consumer Data Research Centre (see page 14)
- ESRC Centre for Doctoral Training in Data Analytics and Society
- Bowel Cancer Intelligence UK (BCI UK)

The LIDA seminar series

The LIDA seminar series provides a regular opportunity for colleagues to come together and hear about the latest developments in and applications of data analytics across a wide range of fields.

The seminars are open to all and we've welcomed over 1200 attendees over the past year, including over 350 colleagues from 52 external partners.

Bowel Cancer Intelligence UK (BCI UK)



Programme lead

Professor Eva Morris

Leeds Institute of Cancer and Pathology

Each year in the UK, over 41,000 people are diagnosed with bowel cancer and 16,000 die from the disease. BCI UK is a research collaboration hosted by the University of Leeds, which works closely with Public Health England, to provide data and intelligence to inform research and care.

Although datasets already exist covering all aspects of care, which could be used to improve cancer outcomes, access for researchers to link and exploit these data is limited. BCI UK is robustly linking multiple electronic data sources to provide intelligence which will underpin patient choice and help individuals reduce their risk of disease and access the best care. Funded by Cancer Research UK, BCI UK is creating a secure data repository called CORECT-R which will contain data from across the cancer pathway of diagnosis, treatment and outcome. This data will be available to all researchers who hold the necessary ethical and regulatory approvals.

www.bci.leeds.ac.uk

EMPOWER

Programme lead:

Professor Susan Grant-Muller

Institute of Transport Studies



EMPOWER sets out to substantially reduce the use of conventionally fuelled vehicles (CFV) in cities through a step change in driver behaviour. Alternatives to individual car use, such as cycling, carpooling, public transit or electric vehicles, already exist in our cities. But to improve urban traffic flows, increase air quality and reduce CO₂ emissions and oil consumption, more people need to use them. Adopting a 'reward rather than punishment' approach, EMPOWER explores the use of positive incentives delivered through smart phone technologies and the web to persuade people to make modest shifts in their transport choices.

Funded by the European Commission's Horizon2020 programme, the project combines empirical research with practical implementation in four Living Lab Cities and seven Take-Up Cities and Communities. Mobility research linked to Empower is also pursued through the Innovate UK funded Catch! Project, the development of the Propensity to Cycle Tool with the Department for Transport, and related work in the ESRC Consumer Data Research Centre and Simulating Urban Flows project.

<http://empowerproject.eu>

QuantiCode

Programme lead:

Professor Roy Ruddle

School of Computing



This cross-disciplinary project aims to develop novel data mining and visualization tools and techniques, to transform people's ability to analyse quantitative and coded longitudinal data.

Whether it is health analysts needing to provide business intelligence for clinical commissioning decisions, researchers modelling disease risk stratification or supermarkets wanting to combine data from purchasing, demographic and other sources to understand consumer behaviour and guide investment, public and private sector organisations require far more powerful analytical tools than are currently available. Today's tools are crude for assessing data quality, frequently involve analysis techniques designed to operate on aggregated, rather than fine-grained, data, and are often laborious to use, which inhibits users from discovering important patterns.

The QuantiCode project addresses these deficiencies by bringing together experts in statistics, modelling, visualization, user evaluation and ethics. The project will deliver proof of concept visual analytic systems, which we will evaluate with a wide variety of stakeholders.

<https://tinyurl.com/QuantiCode>

MRC Medical Bioinformatics Centre

Established: 2014

Funding: Medical Research Council – £7.3m

The MRC Medical Bioinformatics Centre (MBC) aims to create and sustain the infrastructure, facilities, understanding and culture changes to enable ground-breaking and productive bioinformatics research at the interface between the clinic, health records and high volume molecular and phenotypic data sets.

Substantial co-funding from the University of Leeds and shared investment with the Consumer Data Research Centre (CDRC) has been used to create the Integrated Research Campus (IRC) (see page 17). IRC hosts data from the Leeds Care Record, software company TPP, Leeds City Council and the UK Biobank and works with technology partners including IBM, IMS-Quintiles, and Aridhia.

The Centre's work has focused on six areas, developing new insights into the factors that drive disease, and providing indicators that inform treatment, to enable better, more personalised and more effective medicine and healthcare:

- Cancer genomics and malignant melanoma
- Genotype, phenotype and response to radiotherapy in rectal cancer
- Inherited rare diseases
- Proteomic biomarkers
- Patient similarity searches in lymphoma treatment
- Enhancing the value of the MRC-funded rheumatoid arthritis stratified medicine and RA-MAP consortia through linkage to complex datasets

The Centre has generated high impact academic outputs, including 340 linked publications and has leveraged £14.4 million additional funding (source: Gateway to Research, 1st January 2017). Cross Research Council activity has been promoted through, for example, links to the CDRC through the ESRC Obesity Strategic Network (with the School of Geography) and to EPSRC through Qanticode (with the Schools of Computing, Mathematics and Philosophy) and microbubble-enhanced imaging (with the School of Physics).

MBC has taken a leading role in Connected Health Cities which is promoting the IRC as a safe haven or 'ark' for local research partners. The Centre has joined with the Farr Institute, eHIRCs and other MRC Bioinformatics Centres to establish a national bioinformatics infrastructure. The Centre works with academic partners across the country; for example, the Universities of Leicester, Edinburgh, Imperial and University College London are all collaborators on the bowel cancer intelligence hub, coordinated from Leeds and funded by Cancer Research UK (see page 11).

MBC has strong links to the Yorkshire Centre for Health Informatics (YCHI), co-located in LIDA and a long-term innovator in training for health informatics. This partnership has delivered new capacity in MSc provision, and enabled the Leeds undergraduate medical curriculum to include substantial exposure to patient records through collaboration with ResearchOne. Leeds is a partner in the Discovery Medicine Centre for Doctoral Training with Liverpool, Sheffield and Newcastle.

MRC investment has been further leveraged through senior appointments and intellectual capacity through the University of Leeds Academic Fellow (UAF) programme. The Centre has provided Bioinformatics Data Scientist support to the Leeds Omics initiative, a network of 'omics' researchers driven by UAFs from the Faculties of Medicine and Biology.

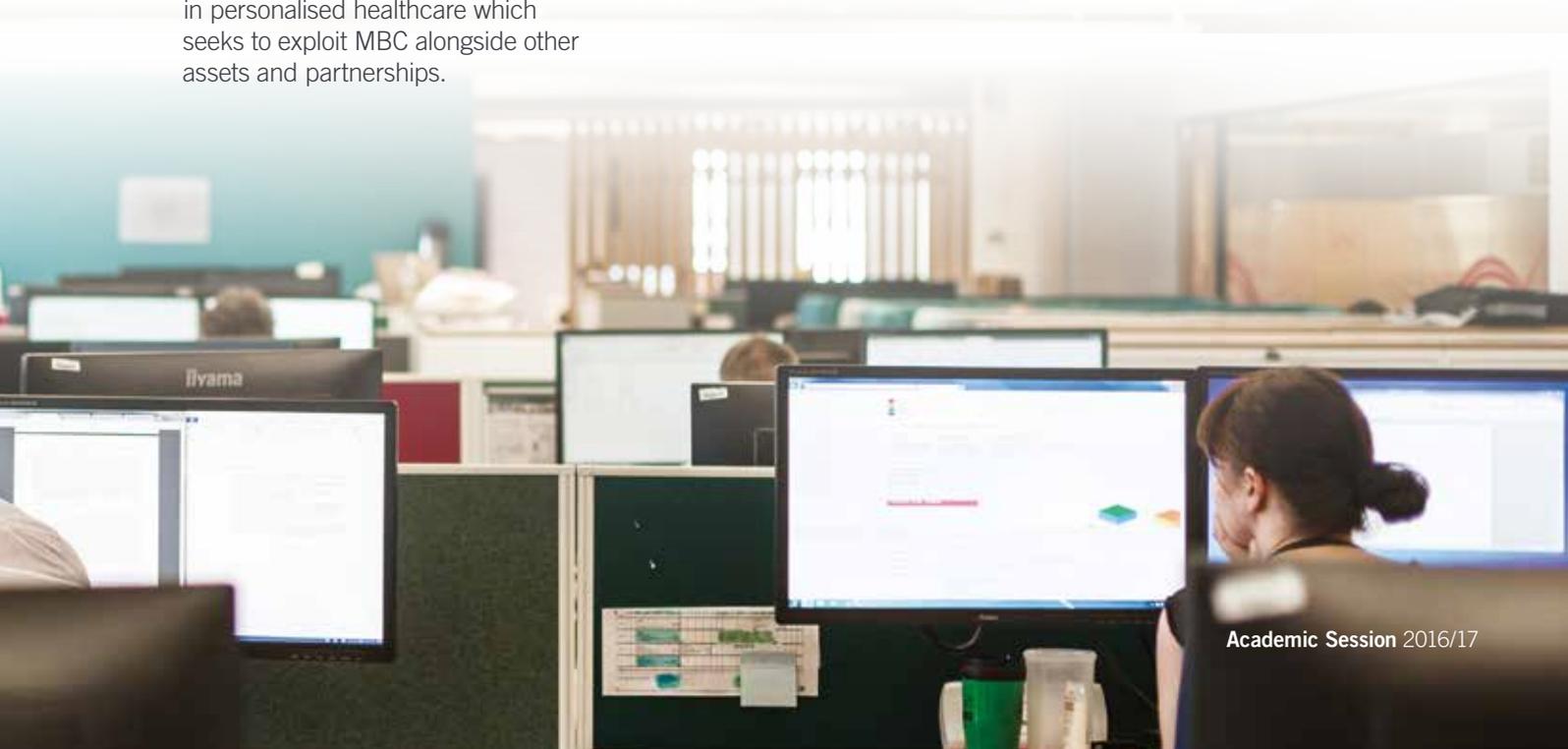
Living in Leeds is a major initiative in personalised healthcare which seeks to exploit MBC alongside other assets and partnerships.

“The Centre has generated high impact academic outputs, including 340 linked publications and has leveraged £14.4 million additional funding.”

A Steering Group, chaired by Professor Sir Alex Markham with support from Geoff Hall (Senior Lecturer in Oncology and Chief Clinical Information Officer in Leeds Teaching Hospitals Trust) and Mark Birkin (Director of LIDA), is planning for the MBC's sustainable future. This will include developing new approaches to personalisation and personalised medicine and working closely with the Leeds Institute for Clinical Trials Research for enhanced impact in eTrials and Applied Health Informatics.

MBC will draw on LIDA's links with mathematics, computing and other foundation disciplines to develop new methods. The Centre will also develop new approaches to research ethics in partnership with patients and the public, such as through the UseMyData project, and through collaborations with colleagues in philosophy and law. Access to distinctive skills and assets such as the CDRC, National Centre for Atmospheric Science and the N8 Policing Research Partnership will provide a unique window on the impact of changes in population health and lifestyle.

lida.leeds.ac.uk/research/mbc





ESRC Consumer Data Research Centre

Established: 2014

Funding: Economic and Social Research Council – £11 million

Led by: University of Leeds, University College London

Partners: Universities of Liverpool and Oxford.

The ESRC Consumer Data Research Centre aims to make data that is routinely collected by business and local government available for research purposes and to undertake innovative programmes of research, training and knowledge exchange.

The Centre has built over 40 partnerships with industry, local government and civil society organisations, making a wide range of data available through a three-tiered data service (open, safeguarded and secure). Our success in securing and maintaining partnerships has resulted largely from our ability to instil trust and to offer focused propositions of mutual benefit. Many of our partners are not only sharing data with us, but also working with us on joint projects, sponsoring interns and PhD students and actively engaging with our MSc programme.

The Centre has over 50 data products from partners including Callcredit, YouGov and Whenfresh/ Zoopla that are accessed via our data portal by individuals from academic institutions, and public and private sector corporations across the country. Safeguarded and controlled data are available following application and approval, while our open data can be accessed via simple download.

To encourage use of our data sets, we are awarding grants from a £500,000 ESRC Innovation Fund to support research projects across a range of disciplines and institutions.

We are also driving forward our own ambitious research agenda and generating impact, in areas such as urban mobility, health and consumer behaviour and ethical and sustainable consumption. One example is the Centre's contribution to the development of the Propensity to Cycle Tool (PCT), which helps predict where cycling has potential to grow and to direct the investment which is needed to promote uptake. This interactive tool has been adopted by local authorities across the country as an important instrument in guiding planning and investment in cycling infrastructure.

Our health and consumer behaviour driver project is increasing understanding of how consumer behaviour is linked to health issues such as obesity, propelled by national and international collaborations through the ESRC Obesity Network.

Our sustainable consumption academics in the Schools of Business and Earth and Environment are helping retailers and local authorities tackle major challenges such as food waste and low recycling rates. Collectively, we are harnessing consumer data and working with our partners for the benefit of society.

The Centre is also investing significant resource in developing data science capability to provide a skilled workforce for the rapidly expanding area of data analytics.

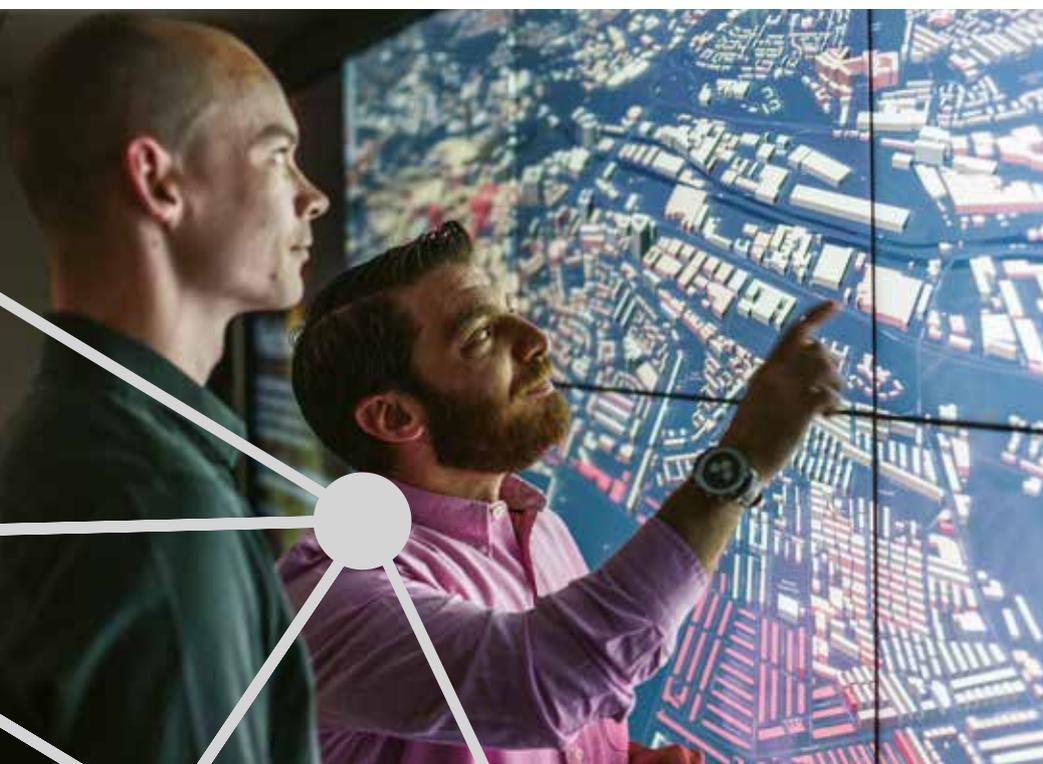
“Our MSc in Consumer Analytics and Marketing Strategy (CAMS) launched in September 2016 with double the projected intake of students.”

The programme is the first of its kind in the UK, offering the novel integration of consumer analytics, spatial analysis and marketing strategy. Students benefit from exposure to real consumer data and direct interaction with business partners. Our business partners are actively shaping the programme, by developing course content, providing data and delivering guest lectures. This year, the programme has grown further and by 2020 it is expected to provide a major funding stream for the CDRC.

We also deliver a wide range of webinars, workshops and short courses and are constantly looking to diversify our training portfolio. Our short courses in R and Tableau and our Big Data focussed training within health, transportation and data visualisation are particularly popular. This year, the Centre launched its first summer school in modelling strategies and published a number of training courses on line.

We aim in the future to offer more bespoke training packages to both academics and non-academics, to help develop the data science leaders of tomorrow.

www.cdrc.ac.uk



LIDA EDUCATION AND TRAINING

LIDA is investing significant resources in developing data science capability, at all levels, in both academia and industry – equipping academics with the skills to deliver impactful research and ensuring that the workforce has the skills to compete in the emerging and rapidly expanding area of data analytics.

Masters Courses

Masters courses in data science provide graduates with high-level general skills and advanced specialist knowledge based on cutting-edge research, in preparation for a career or further study. LIDA academics teach and supervise students on courses across the University, including:

- MSc Consumer Analytics and Marketing Strategy (CAMS)
- MSc Advanced Computer Science (Data Analytics)
- MSc Data Science and Analytics
- MSc Epidemiology and Biostatistics
- MSc Business Analytics and Decision Sciences
- MSc Geographical Information Systems
- MSc Health Informatics
- MSc Health Data Analytics (2018).

LIDA also enables students to experience working with real life systems and data, with medical undergraduates using the TPP SystmOne clinical information system (www.tpp-uk.com) and CAMS students working with consumer data via the CDRC.

Data Scientist Internship Programme

Launched in September 2016, this programme offers graduates the opportunity to undertake a 24-week internship at LIDA working with multidisciplinary research teams across each of LIDA's constituent areas: medical bioinformatics, consumer data research and multidisciplinary data analytics.

The programme enables interns to work alongside leading scholars, own delivery of a project, get hands-on technical experience using real data and establish links with external project partners.

In its first year, the programme planned to offer eight internships, but the number and quality of applicants was so high, that in the end 10 graduates were appointed.

They completed the programme as a cohort, building skills and knowledge in advanced analytics, through mentorship, applied research, and through on-site training opportunities in statistical analysis, visualisation, research methods and computer programming.

The majority of our 2016 interns have already secured PhD studentships or employment in data science roles. Our next cohort of graduates will start their internships in September 2017.



Centre for Doctoral Training in Data Analytics and Society

Funded for four years, the ESRC Centre for Doctoral Training will provide postgraduate research and training to 52 PhD students across the Universities of Leeds, Liverpool, Manchester and Sheffield.

Launched in 2017 and led by the team at LIDA, the Centre – a collaboration with external partners – provides the opportunity and funding for graduates to undertake a four-year integrated PhD and MSc in Data Analytics and Society.

The programme covers core research training and domain skills for data science, combined with an internship and research project working with external non-academic partners, such as Shop Direct, YouGov, South Yorkshire Police and Costain.

INTEGRATED RESEARCH CAMPUS

University Academic Fellows

Two hundred and fifty exceptional individuals with the potential to make a major contribution to the University's academic performance are to be recruited to Leeds as University Academic Fellows. Thirty-six of the appointments to date are aligned to LIDA.

While the appointments align to core disciplinary strengths and strategies, they are typically based in strong cross-disciplinary activity or emerging academic platforms and support the major external opportunities and themes for growth identified by the University.

Training and Capacity Building

LIDA delivers an annual programme of training for both academics and non-academic researchers. The programme comprises introductory courses for postgraduate students through to advanced training for data scientists, including:

- Introductory and advanced R courses
- GIS
- Spatial methods
- Visualisation and Tableau
- Predictive Analytics
- Spatial methods for public health researchers
- Advanced modelling strategies

Find out more at

lida.leeds.ac.uk/study-training

The Integrated Research Campus (IRC) supports large scale research through infrastructure, training and services for secure data handling.

The IRC Data Services Team, made up of data-focussed analysts and developers, are managed in collaboration with University IT Services, but located in LIDA, providing shared services and working with partners across the University and the city. The IRC Data Services Team has a mix of skills in systems and database administration, data management, software engineering, data analysis and data science and information governance expertise. They support researchers at all stages of the data-intensive research lifecycle and are responsible for information security and classification, risk assessment and ensuring compliance within legal and ethical frameworks.

IRC services include dynamic data capture, processing, analysis and visualisation, with all data handling subject to relevant agreements, approval or consent. IRC infrastructure includes web services and gateway applications, visualisation suites, safe rooms and a Virtual Research Environment (VRE), where researchers can collaborate securely within a tailored and scalable, firewall-protected research platform.

The IRC achieved accredited certification in 2017 to the international standard for information security management, ISO/IEC 27001:2013.

By implementing the standard, the IRC established a robust approach to managing confidential information, which is integral to its commitment to protect the confidentiality, integrity and availability of information and information systems.

The IRC has also been working towards compliance with information governance requirements set by NHS Digital and the Department of Health, known as the NHS IG Toolkit (IGTK). The IRC's information management system has been reviewed up to the minimum 'level 2' and is deemed to be satisfactory. The IRC is therefore authorised to facilitate the NHS Digital Data Access Request Service and support researchers when their requirements include the processing of personal and/or sensitive data (including health and social care data).

Compliance work in 2017/18 will focus on maintaining ISO accredited certification and evidencing NHS IG Toolkit Level 3 compliance, through spot-checks and regular monitoring of IRC controls, including working with an external audit partner as part of the IRC's wider assurance framework.

LIDA PARTNERSHIPS



We believe in innovation through collaboration and work closely with a wide range of data partners in the public, private and third sectors to ensure the Institute's academic expertise is channelled towards addressing real-world challenges.

Our partners include: AQL; CallCredit Information Group; Department for Transport (UK); the Digital and Transport Systems Catapults; IBM Watson; Leeds City Council ; NHS Digital; Public Health England; Zoopla/Whenfresh.

Partners can engage with LIDA through:

- Co-location of their staff within the institute, either permanently or for a defined period relating to a particular project or activity. Whilst working at the Institute, colleagues benefit from direct access to members of the academic and technical teams, whilst also making use of our facilities, including the Integrated Research Campus (see page 17).
- Collaborative research, from short proof-of-concept projects spanning a few weeks or months to long-term programmes over several years. Our flexible and adaptable approach allows us to structure projects and programmes to suit the needs of our partners.
- Our Data Scientist Intern programme to test new ideas, methods and techniques before committing to a longer term programme of work. As an Intern project sponsor, our partners set the project brief and then receive regular updates and a final report from the intern and their academic supervisors.
- Sponsorship of PhD projects for longer term research, enabling partners to access skills, strategic advice, data and networks. As a PhD project sponsor, our partners work with the academic supervisors to set the initial brief and remain engaged to help steer and inform the project throughout its duration.
- Guest lectures and masterclasses cover both real world applications of data analytics and related career opportunities and enable partners to help shape the skills of the future workforce.
- Training and professional development. Our programme of short courses are open to our partners and we can also meet bespoke training requirements.

Callcredit Information Group

Callcredit has a longstanding relationship with the University of Leeds, providing its data sets and marketing systems, such as the CAMEO consumer insight products, for use in research and postgraduate teaching. The partnership with LIDA is enabling more in-depth and detailed geodemographic data – and advice on how best to use it – to be shared with the academic community via the Consumer Data Research Centre for groundbreaking research proposals, with Callcredit gaining insight into emerging new tools and methods for data handling and modelling.

The collaboration has already secured funding for a three-year PhD project, to develop a new geodemographic profile of Leeds, working with the City Council. The classification will help the Council tackle challenges around service provision, policy development and understanding health in the city, enabling them to deliver more tailored services.

Potential applications include identifying segments of the population most likely to have certain health outcomes, to ensure appropriate locations for healthcare services, or understanding how people move around the city, to provide essential information for public transport providers and infrastructure planners.

LEEDS CITY COUNCIL

The Intelligence function of the Council has been working with LIDA since its foundation in July 2015. These Council teams have a permanent base at LIDA enabling them to network with university staff, access powerful IT resources and work closely with academic staff on specific projects. The Council is using this partnership to:

- develop tools to better understand changes in an individual's social care needs over time to enable the Council to proactively manage their care, and ensure they stay in their own home for longer
- develop a customer segmentation model for Leeds
- bid for funding to develop a model for testing how efficiently health and social care resources in the city are allocated
- better understand flooding patterns in Leeds following the 2015 Boxing Day floods, and help better manage a similar event in future
- better understand the differing level of resilience across different Leeds communities
- access training in statistical and analytical techniques and software packages.



Alternative formats

If you require any of the information contained in this publication in an alternative format e.g. Braille, large print or audio, please email [**disability@leeds.ac.uk**](mailto:disability@leeds.ac.uk)



UNIVERSITY OF LEEDS

University of Leeds
Leeds, United Kingdom
LS2 9JT
Tel. +44 (0) 113 243 1751
www.leeds.ac.uk