Institution: University of Sheffield

Unit of Assessment: 11 - Computer Science and Informatics

a. Overview

The Department of Computer Science's submission consists of 31 staff, 56 RAs and 97 PhD students. We structure ourselves into seven research groups: Computational Biology (CompBio); Machine Learning (ML); Natural Language Processing (NLP); Organisations, Information and Knowledge (OAK); Speech and Hearing (SpandH); Verification and Testing (VT) and Virtual Reality and Graphics (VRGraphics). Staff can of course belong to more than one group.

Interdisciplinary research is a strategic focus at Sheffield, and this is strongly reflected in our activities: since 2008 we have been founding members of 4 new interdisciplinary research centres:

- SITraN: *Sheffield Institute for Translational Neuroscience*
- INSIGNEO: the *Institute for in Silico Medicine*
- CATCH: the *Centre for Assistive Technology and Connected Healthcare*
- SCentRo**:** *Sheffield Centre for Robotics*

We have extensive collaborations within the Engineering Faculty and across the University.

b. Research strategy

Our vision is to achieve research excellence, innovation and impact, as an internationally leading Computer Science department, translating fundamental knowledge into practice, through collaboration with end users. We **aim** to achieve impact by transcending the usual boundaries between engineering, medicine and the life sciences. From a methodological point of view two themes underpin our research groups: **extraction of knowledge from data** (NLP, OAK, SpandH, ML, CompBio) and **modelling and simulation of natural and computational systems** (CompBio, VT, SpandH, VRGraphics).

Our approach. We invest in core computer science research, advancing theory and practice. Each research group has a *scientific* focus, and an *engineering* focus. The **scientific** focus contributes to the fundamentals of the discipline, in the long term enabling large steps forward in technology. For example, our work in theoretical CS work lies not just in VT, but also in ML and CompBio. The **engineering** focus ensures that our contributions have a direct impact on practice. We deliver software and applications directly to our collaborators and end users. E.g., our work on GATE (see Impact case study) and on clinical and assistive technology (see SpandH below). We **collaborate** across the University as a founding member of new interdisciplinary research centres (SITraN, INSIGNEO, CATCH, SCentRo). We collaborate with industrial and clinical partners (74 partners on current projects) to deliver real **impact**, and our new *Advanced Computing Research Centre* (ACRC) will strength collaborations with industry (see impact documents).

Progress against RAE 2008 plans. As planned, our strategic investments have focused on expanding the department's expertise in **CompBio** and **ML** and sustaining our strengths in **software engineering**, **speech** and **natural language processing**. Since 2008 our strategy has evolved to include strategic investment in **personalised health** as a logical extension of our computational biology work, and the creation of the **OAK** group to further the impact of our knowledge management work. As planned, the strategic importance of ML has been reinforced by 3 new appointments each of whom works at the interface ML and NLP, SpandH or Comp. Biology.

We have enhanced our profile substantially since 2008, with growth in all relevant metrics: 12 new research active staff, 3 new chairs, 6 new fellowships, new research awards up to £6M p.a. (now FTE £194k p.a. – almost double that of 2008), increased citation rate, 5 new spin-outs etc.

The **Natural Language Processing** and **Organisations, Information and Knowledge** groups (7 staff, 29 RAs, 25 PhDs) are internationally leading groups (e.g., 10 papers at ACL'13) working on the automatic processing of languages, and on the understanding of large scale distributed information. NLP's scientific focus is on developing corpus-evidenced computational models of *understanding*, *translation* and *generation* of language. Achievements include work on the semantics of time and events, word sense disambiguation (*Stevenson's* EPSRC Fellowship), and groundbreaking work on Bayesian non-parametric parsing. Its engineering focus is on applications of language technology in, e.g., document summarisation (*Bontcheva's* EPSRC Fellowship), machine translation, and information extraction, c.f. the GATE framework (see Impact

Case Study).

Since 2008 increasing work in **NLP** on knowledge management and its application to industry led to the creation of the **OAK** group, with an engineering focus on developing methodologies for assimilating information from multiple distributed devices and delivering actionable knowledge. Applications have included a large knowledge management system for Rolls-Royce and a web-scale system for monitoring social media during periods of civil unrest (see Impact Case Study).

SpandH (4 staff, 10 RAs, 18 PhDs) studies computational models of speech communication as a means to build robustly engineered speech technology. Since 2008 this strategy has enabled the group to broaden its range of application areas and to expand its reputation for building large-scale speech systems. Conversational speech transcription research developed into the pursuit of "Natural" Speech Technology (NST), a grand challenge supported by a £6.2M EPSRC programme grant (Sheffield, Cambridge and Edinburgh). This work attracted industrial support (Cisco, Google) and led to spin-out companies (koemei, SA; Quorate Technology). Work on computational hearing is now underpinning applications in speech enhancement and hearing aid signal processing.

In **Verification and Testing** (7 staff, 7 RAs, 27 PhDs) work on specification and verification provides a strong *scientific* focus. Achievements include algebraic computational models, foundations of concurrency and its verification, and exploiting state-based refinement methodologies. Mechanised proof is a feature of this work. Generalised models of membrane systems have seen application in synthetic biology. Our *engineering* focus lies in testing, where strands on model-based, property-based and evolutionary testing have produced toolsets such as EvoSuite, StateChum & IGUANA. We have, e.g., provided the first real empirical evidence for the value of search-based testing. EPSRC and EU funded work on model-based testing and reverse engineering has found industrial application.

Since 2008 our investment in **Computational Biology** (6 staff (incl. 3 ECRs), 4 RAs, 8 PhDs) has led to new computational approaches to study structure and function in biological systems, with an increasing emphasis on integrative models for personalised healthcare. Scientific research has included work on evolution and behaviour, with a particular focus on animal groups and computational neuroscience. The INSIGNEO institute for in-silico medicine exemplifies our engineering focus with work on novel computational approaches for the study of heart disease, tissue engineering and aneurysm formation. *Clayton* is leading a £1.52m EPSRC project on multiscale modeling and *Hensman* has an MRC Fellowship for high throughput genomic data.

Work on machine learning spans many groups, with the **Machine Learning** group (4 staff, 5 RAs, 10 PhDs) focusing on developing machine learning *methodologies*: it develops new approaches to data analysis inspired by challenges in real applications, underpinning many of the achievements in language, neuroscience, health and genomics. The group has an excellent international reputation (2014 NIPS Program Chair, IEEE PAMI Ass. Ed. In Chief).

In **VRGraphics** (3 staff, 1 RA, 9 PhDs) our scientific focus is on producing highly realistic simulations of natural phenomena (e.g. of touch, human interaction, facial movement). This is applied to engineering problems through simulations using virtual environments.

Strategic plans for 2014 onwards. By being part of a large and financially buoyant Engineering Faculty (e.g., now 2nd highest Eng. research income in UK) we have been able to invest heavily in research. Our own turnover has increased by 27% to £12M. We have created reserves of £1.3M for future strategic investment in the next period, and have had access to £6.4M of Faculty resources to support strategic initiatives, e.g., start-up costs of new senior staff. We will continue to exploit the opportunities this brings, by, e.g., expansion of research space within the new building programme. The Faculty has committed to a significant expansion of the department, with further staff appointments at both senior and junior level. This will allow us to realize our **future strategic research directions** that span our research groups, structured around three emerging themes, and for which the **ACRC** will provide a route to industrial exploitation:

- **Extraction of knowledge from data.** Massive volumes of data, particularly unstructured data, are being generated by sensors / organisations / individuals, raising huge challenges of interpretation, aggregation and summarisation. We will expand our globally leading expertise in ML, speech and language to work on areas such as speech recognition in noisy environments, interpreting social media, and in new areas such as mining conversational data, and discovering knowledge in biomedical data. Methods for dealing

with unstructured data will be a focus for our work.

- Our **modelling and simulation** theme addresses the importance of complex systems analysis. We will enhance our position as one of the largest UK groups on testing, developing new model- and search-based approaches. We expect to see a significant growth in work on algorithms – their analysis, verification and simulation, exploiting expertise across VT, CompBio and VRGraphics.
- **Personalisation** has the potential to revolutionise the role of computers in our lives. We are uniquely placed to work on a range of areas such as personalised health, speech recognition, and personalised information assistants. Work in this theme will develop interfaces and methods for tailoring the individual experience, and will link to our work in computational biology and health.

We will exploit centres such as **INSIGNEO** and **SITraN,** so that work in **CompBio** and **ML** has a major impact in the next period, and we will invest in new appointments in the interface in these groups and **INSIGNEO**, **CATCH** and **SITraN**. We will see increasing use of computational models to influence clinical practice in diagnosis and intervention, and see work in **VRGraphics** increasingly being applied in computational biology and health.

c. People:

i. Staffing strategy and staff development

Our ***Strategy*** for academic appointments reflects our interest in both core computer science and its application. We are committed to appointing and retaining leading international researchers at all stages of their careers. Our investments (12 new staff) were focused on expanding our expertise in **CompBio** and health whilst sustaining it in software engineering, **ML** and **NLP**.

In 2009 we appointed *Vasilaki* (computational neuroscience) and *Cohn* (probabilistic natural language modeling) at Lecturer level. In 2011 we made three further lecturer appointments: *Sudholt* (CompBio / VT), *Fraser* (testing, recruited from Saarbrucken) and *Specia* (NLP). In 2013 *Watton* (computational BioMedicalEng) joined from Oxford as a core member of INSIGNEO.

At a more senior level we have recruited *Lawrence* (ML and CompBio, 2010, from Manchester), *Marshall* (CompBio, 2010, from Bristol), and *Harders* (Haptics, 2013 from ETH Zurich). *Lawrence* has a joint chair with the Sheffield Institute of Translational Neuroscience (SITraN). *Harders* strengthens our links with INSIGNEO and the Centre for Assistive Technology and Digital Healthcare (CATCH). *Marshall* was appointed on a prestigious Prize Readership and strengthens our interdisciplinary portfolio. We are in the process of recruiting an additional Chair to work explicitly with the CATCH centre. Three new fellows (*Hensman*, *Oliveto* and *Richmond* – see below) will progress to lecturing posts at the end of their fellowships.

Support for new staff: All new appointments are given the opportunity to establish themselves within the department with an initial reduced teaching and administrative load for their first three years in post. Particular support is given to new lecturers. Specifically, we offer:

- a half teaching load during probation,
- no admin duties for 3 years,
- start-up travel and equipment funds (£10k),
- a fully funded PhD scholarship,
- preferential treatment in obtaining fee scholarships for overseas PhD students.

To support time on new funded projects (e.g. EPSRC first grants) we double the amount of allocated time in our workload allocation model (e.g. raising 10% to 20%) for grants for probationary staff. In addition, we commit extra PhD studentships to all EPSRC grants. Our support is taken as a model across the University, and is widely praised, e.g. all EPSRC reviewers for new staff's first grant have commented positively, and one wrote "*the support offered by the HoD and dept is outstanding and should be a model for every University"*.

Our approach has been highly successful. *McMinn* (Testing, appointed prior to 2008, promoted to SL during assessment period) has an excellent profile of outputs and 5 EPSRC grants. *Vasilaki* and *Cohn* both recently completed their probationary period with excellent track records of funding (EPSRC, EU, industrial grants) and publications (see outputs). *Cohn* has just been awarded a 5-year EPSRC Early Career fellowship. Already, within two years of appointment *Specia* has £805k grants, *Fraser* has secured a major grant from Google, an EPSRC grant, and a Microsoft award,

and *Sudholt* a FET Young Explorers grant.

Our recruitment has been funded by significant Faculty and University investment: *Marshall*, *Oliveto* and *Richmond* from Fellowship schemes, plus headroom from staff movement. Retirements included *Green* and *Holcombe*, who still play valuable roles, e.g., *Holcombe* continues to represent the University and will direct the ACRC (see Impact template). *Green* continues to play a leading role in the EPSRC Natural Speech Technology Programme Grant and the SpandH group.

Career development We have a wide range of formal and informal mechanisms to support staff during their career. The annual SRDS scheme provides an appraisal and self-reflection opportunity: it assesses each year's performance, rewards exceptional contribution, and importantly it includes a goal setting and development component. We have a formal *Workload Allocation Model* which records measurable activity and effort under three headings: Research, Teaching and Admin., to prevent individuals being overloaded. This allows reduced loads to be given to our star performers in research, and we also support academics targeting particular calls for prestigious grants. Part-time staff benefit from sabbaticals and pro-rata loads.

Every new member of academic staff is assigned a personal mentor to provide further support and guidance. The Head of Department (HoD) also meets with new academic staff informally every month to provide an opportunity to talk through any research, teaching or University issue that may be pertinent to their careers. The University provides training courses for new staff, in topics such as "proposal writing", "research supervision", and "research project management". In addition to a reduced load for new staff, we support staff in post through a planned programme of sabbaticals (1 per 7 years), and we commit to support two full sabbaticals per year across the department (12 since 2008).

The wealth of senior staff has enabled us to identify and develop individuals who will become future research group heads and senior managers, and promoted accordingly to deal with renewal. The University runs a Sheffield Leaders programme, specifically designed to equip individuals for significant leadership roles (*Clayton*, *Gaizauskas, Stevenson*, *Barker*, *Marshall* have taken part). We have identified future research leaders in the department and promoted accordingly to deal with renewal. Since 2008 *Barker*, *Romano*, *Stevenson*, *McMinn* were promoted to SL, *Struth* and *Gheorghe* were promoted to Reader, and *Hain* and *Brown* to Chairs.

Career development of researchers We are committed to applying the principles of the *Concordat for the Career Development of Researcher*s. We are in the process of moving research staff from fixed-term to open-ended contracts. We use bridging funds to keep researchers in employment if there is a gap between funded proposals. To help with their career progression we run activities such as grant and fellowship writing workshops, and run specific career guidance sessions at our annual Research Away Day.

Support at key transition points is available. E.g., a *Mentoring for Research Staff* scheme is run for researchers and a *Researcher Mentoring Manager* has been appointed. Our *Eng. Researcher Society* connects and supports PDRAs with a programme of events. We encourage senior RAs to apply for promotion to grade 8 (i.e., L scale), allowing them to develop their career in research (6 promoted in period, including 3 female). Researchers who are awarded a prestigious fellowship are automatically considered for an open-ended Lecturer position at the end of the funding period.

Fellowships An extended period of research focus adds depth to staff careers, and we strongly support our staff in seeking fellowships. *Bontcheva* started an EPSRC Career Acceleration Fellowship in 2010, and will join the academic staff on its completion; *Stevenson* completed an EPSRC Advanced Fellowship in 2011. *Cohn* has just been awarded a 5-year EPSRC Advanced Research Fellowship. *Hensman* began a four year MRC Special Training Fellowship in 2013, collaborating with SITraN. In addition, *Romano* completed a RAEng Industrial Fellowship for her work with COSTAIN. *Sharkey* has been making major contributions in public awareness and understanding of science and was funded to do so by Leverhulme and EPSRC Media Fellowships (see Impact case study). *Walker* was an RCUK fellow to 2009 before joining the academic staff.

Marshall was appointed via a Faculty Prize Fellowship, and is supported via associated RAs and PhD studentships. A University-wide VC fellowship scheme led to the appointment of two new Fellows (out of just 16 in the University): *Richmond* in VRGraphics and *Oliveto* in algorithms, both of whom will progress to lectureships at the end of the fellowship.

International dimension and visitors Our international profile is reflected in the nationalities of recent recruits. Our 12 new staff since 2008 include six different nationalities, arriving from many leading overseas institutions (Saarbrucken, ETH, EPFL, etc). This is mirrored in the diversity of our researchers. We have similar mobility at PGR level, supported in part by our involvement in several ITN Marie Curie schemes. We have a regular stream of prominent visitors to the department, e.g., *David Harel*, *Geoff Hinton*, *Peter O'Hearn, Simon Tavare*, in our regular seminar programme.

Equality and diversity The University's *strategy* and *action plan* provide mentoring opportunities for junior and senior staff. Within the Faculty, *Derrick* (HoD) has led work on *equality and diversity*. This has resulted in the appointment of a Faculty Director of Women in Engineering together with a team to implement activities to support women at all levels. We provide career development workshops, sessions on career progression, leadership skills, mentoring opportunities, and have set up a network for women in STEM. We have provided media training to remove barriers that may prevent women from speaking confidently. We have a very strong supportive culture of allowing flexible working, which many take up. We help female academic staff returning from maternity leave with the transition, e.g. in the last 3 years 3 CS academics have received funds from the *Women Academic Returners Programme*. We hold an *Athena Swan Bronze* award in the dept.

We have provided research-only career paths for some of our best female researchers (e.g. *Bontcheva* PI on several major EU grants, now an EPSRC fellow, *Walker* until recently an RCUK fellow), and we have more female researchers at G8/9 (L/SL grade) than males. At L/SL level we now have 40% female staff, and expect to have several promotions to R/Prof in the next few years. We ensure females are well represented on all committees and decision making bodies.

ii. Research students

Since 2008 we have devoted considerable resources to improving the research student experience. The number of PhDs awarded has grown from 13p.a. to 18p.a. and intake has increased by 50%. We are committed to **recruiting** the highest quality students. Our best applicants are put forward for Faculty/University scholarships, and our department budget surplus (since 2008 £500k spend: 18 students) augments our DTA provision (at least 2 per year). Around 10% of our PhD students now have industrial financial support.

We have led and participated in a large number of EU Marie-Curie training networks (e.g., S2S, SCALE, INSPIRE, MLPM). These not only fund PhD students (and also ESRs that aren't included in the count of PhD students as they are on staff contracts) but have also funded lengthy PhD internships and networking activities such as locally-hosted summer schools and workshops that increase the dynamism of the PhD experience across whole research groups within the dept.

Our overseas students come with external funding from e.g., the Saudi Arabia's *KASP*, Brazilian *Science without Borders* and Mexico's *CONACYT* programmes. We have worked with sponsors to ensure that we can meet their expectations in terms of, e.g., progression and reporting. These changes have led to a steady stream of high quality fully-funded overseas students.

PhD Training and support mechanisms: Each PhD is given personal lab space, a new PC and access to our research network and the White Rose HPC cluster. Since 2009 we invested £62k in PhD equipment and since 2011 £38k for PhD travel: each student receives up to £5k travel funding. We encourage excellence by prioritising top conferences and journal publication costs. The Faculty *Learned Societies Fund* has provided a further £30k since 2010 for staff and students.

Support and Monitoring: Students have weekly meetings with their supervisor, and are supported by a three-member panel that meets every 6 months and reviews research progress to ensure a smooth path towards timely completion. This process allows for monitoring a student's progress in the longer term and assesses the criteria for progression at a 12 month meeting. We have significantly improved our 4-year completion rate since 2008, which is now the best in the Faculty.

Training: Our *Doctoral Development Programme* provides flexible student-centred training, driven by a regular training-needs analysis. A student's programme is built from a broad range of modules covering academic/technical skills, presentation and communication, career development etc., provided across the University. The University also recently initiated a teaching skills course that PhD students take as a pre-requisite for laboratory demonstration.

We benefit from being part of a Faculty that is committed to delivering excellent inter-disciplinary research training. By 2014 the new *Engineering Graduate School* will provide state-of-the-art

research facilities for robotics, virtual environments and interdisciplinary centres allowing us to continue to attract top students.

Additional networks (such as Marie-Curie ITNs, or the Midlands Graduate School in the Foundations of CS) provide further training opportunities in subject specific areas. We have a regular stream of students undertaking (industrial and academic) visits and internships (e.g. at Microsoft). We participate in the organisation of prestigious summer schools (Johns Hopkins CSLP Workshop, ML Summer Schools etc).

Destinations: Many of our PhD students go on to academic positions and research posts. Recent graduates have left for Google in Mountain View (Guthrie, 2011), Microsoft Research, LA (Fusi, 2013), Red Hat (Hughes), in addition to many academic posts.

d. Income, infrastructure and facilities

Income: Our funding base has grown year on year. Since 2008 we have won over £18M. In the last 3 years alone we won over £15M of research funding. Large recent grants include the NST programme grant in SpandH; AirProm, Greenbrain, Quintet (EPSRC) in CompBio; and Arcomem, WeKnowIt and WeSenseIT (all EU) in NLP and OAK.

	2008-9	2009-10	2010-11	2011-12	2012-13
New grants awarded	2,285k	1,625k	5,270k	4,130k	6,000k
£ Per FTE (31 staff)	109k	71k	188k	148k	194k
Consultancies	7k	6k	94k	14k	18k

Our funding comes from over 40 sources, and its diversity reflects the interdisciplinary nature of much of our research. This includes research councils (EPSRC, BBSRC, AHRC, ESRC, NERC, ERC), charities (Arthritis Research Campaign, Leverhulme, RNID, Royal Society, Wellcome, Yorks Cancer), government (DTI, GCHQ, Human Frontier Science prog, JISC, MoD, TSB), large companies (BAE, Cisco, DSTL, Google, IBM, Microsoft, Proctor & Gamble, Rolls-Royce) and SMEs (e.g. Blitz Games, ITSLanguage, Matrixware). The depth and breadth of our funding is reflected in our portfolio: each group has new EPSRC, EU and either industrial or consultancy funding in the REF period. Apportioned research income by research group during period:

	CompBio (6 pers)	ML (4 pers)	NLP / OAK (7 pers)	SpandH (4 pers)	VT (7 pers)	VRGraphics (3 pers)
Charity	267k	179k	10k	61k	21k	75k
EU	17k	526k	7,224k	942k	1,152k	54k
Govt			293k	380k	21k	151k
Industry	251k		700k	238k	184k	
RCUK	1,178k	747k	2,766k	2,341k	2,131k	134k
Total	**1,713k**	**1,453k**	**10,994k**	**3,962k**	**3,668k**	**414k**

In **CompBio** EU funding supports the development of healthcare infrastructure (*Walker*), charity (*Clayton*, *Walker*) and industry support model evaluation, and EPSRC (*Clayton*, *Marshall*, *Walker*) the development of novel models. BBSRC (*Marshall*) funds basic science research in bio-inspired robotics. *Walker* is Sheffield PI on a new IP (CHIC) about to commence. **ML** has been supported by the EPSRC, MRC and EU for applications (*Cohn*, *Lawrence*) and high risk research (pump priming e-futures), by EU for theory and charities for blue sky research (*Vasilaki*).

In **NLP** development of the GATE platform (*Bontcheva, Cunningham*) has been supported by the EU, EPSRC, JISC, IBM (x17 grants, incl. 3 EU co-ordinations). Work on biomedical text analysis (*Gaizauskas, Hepple, Stevenson*) has been supported by the MRC and EPSRC (x3); on machine translation (*Cohn, Gaizauskas, Specia*) by the EPSRC, EU and Google (x6); on word senses (*Stevenson*) by the EPSRC (x2); on social media analysis (*Bontcheva, Cohn, Gaizauskas, Hepple*) by the EU and EPSRC (x3); on vision and language (*Gaizauskas*) by CHIST-ERA.

In **OAK**, *Ciravegna's* team of 7 RAs has been funded by over 10 grants including coordination of both EU €13.8M XMedia (€2.4M for Sheffield) and EU €6.5M WeSenseIt (€1.1M for Sheffield); and funding from Rolls-Royce worth £390K over 3 projects.

In **SpandH** AHRC, EPSRC and RNID support computational hearing (*Barker*, *Brown*); NIHR clinical speech technology (*Green*, *Moore*); "natural" speech technology (*Hain*) is supported by industrial funding (Google, Cisco) and a £6.2M EPSRC programme grant (2011-2016). The group

has also participated in three Marie-Curie ITNs (S2S, SCALE and INSPIRE).

In **VT** highlights include extensive support for testing: 2 EU grants (*Derrick*, *Bogdanov*: ProTest, Prowess), 5 EPSRC (*McMinn(2)*, *Bogdanov(2)*, *Fraser*) plus industrial support from Google, Microsoft (*Fraser*). Work on verification includes two EPSRC grants (*Derrick*, *Struth*), and that on modelling and FLAME includes EPSRC, BBSRC, NERC (*Gheorghe*) and BAe (*Derrick*). In **VRGraphics** *Romano's* work on agent-based simulation has been supported by the EPSRC, DSTL and COSTAIN (via RAEng). *Harders* has already secured an EU project, Beaming.

We have an outstanding record for leadership in collaborative research, particularly in **EU programmes**. Sheffield is particularly well resourced to lead large EU projects, with a dedicated Framework Programme Office. Since 2008 we have participated in 33 EU projects: 21 Collaborative Projects (8 IPs / 13 STREPS) and 12 others (CSA's, ITN's). Of these we have coordinated **7** Collaborative Projects (AnnoMarket, Arcomem, Companions, Paths, ProTest, Prowess, WeSenseIt) and played a leading part in 3 more (ForgetIT, Khresmoi, Seals). The GATE team within NLP has had a remarkable run of success, bringing in **8** projects (coordinating **3** of them). Its **12** members are entirely self-supporting.

The EU (£9.9M new awards) and RCUK (EPSRC new grants £8.3M, other RC £1.1M) have been strategically important sources of funding since 2008, and are expected to remain so in the future.

Our EU collaborations also allow us to not only interact with the leading research groups in Europe, but also engage with the user base of our research in very dynamic ways. Industrial funding and consultancies include pharmaceuticals in CompBio, a leading Formula 1 team in ML, Matrixware in NLP, Google in NLP, SpandH and VT and Rolls-Royce in OAK.

Looking forward we remain committed to engaging in EPSRC and EU, as many of our research groups align with their priority areas. VT are well placed to exploit growth in verification and correctness (EPSRC shaping capability). *In-silico* medicine is likely to be a high priority in Horizon 2020, and INSIGNEO will also pursue EPSRC support in the core engineering research (such as our recent EPSRC Frontier grant on multiscale modelling), and are well placed to exploit funding from, e.g. British Heart Foundation, Wellcome, MRC etc. However, our diversity of funding ensures that we are not solely reliant on, e.g., EPSRC. We expect industrial funding to increase as we exploit and develop strategically important relationships, e.g., COSTAIN and Rolls-Royce (who spend ~£7M p.a. across the Faculty). The ACRC will play a major role in ensuring sustained industrial support, going forward.

Our physical **infrastructure** is based around a purpose-built Laboratory housing offices and specialist labs for individual research groups. The creation of the **faculty** structure in 2008 has enabled a coherent approach to research and innovation strategy that has led to significant investment, growth and success. A flagship **building programme** includes a new £21M *Engineering Graduate* building (opens 2013, 5.4k m^2) that will house INSIGNEO and SCentRo. The SITraN institute has new purpose-built accommodation which houses a portion of our CompBio work. Construction has just begun on a new £81M *Engineering Building* (19.5k m^2) that will release space in the department to house further research labs and offices. This new investment will provide a focus for all lab work on CompBio (currently in the Kroto building), Robotics (which underlies much of our work on speech) and ML. The Univ. is also investing £53M from 2012-2017 in a **major refurbishment** of 12k m^2 of the Faculty's estate to accommodate the growth in academic staff and associated laboratory space, having already spent £21M since 2008.

Facilities: Our department **Research Office** comprises of two admin staff and two secretaries, providing both pre- and post-award support. Individual groups have additional admin support that work in collaboration with this office. Our admin staff work closely with the University Research Office, Finance and HR. In addition, the Faculty's research infrastructure provides support for large grant applications, and manages strategically important relationships. We provide departmental funds (since 2008: £463k) to research groups and individuals to support early stage collaborations and to pump-prime activity, and we pioneered *red-lining* weeks for grant writing where all other activity is suspended so that academic staff can focus on new research plans.

Support over and above central IT resources is provided by a **Technical Support team** of six in the department. Specific support includes a purpose-built server room for hosting compute and storage machines. By default grant holders are provided with virtual machines within this facility for hosting project web pages and code repositories.

Different groups have specific computational and storage needs and labs have been equipped using strategic reserves, e.g., SpandH has an instrumented meeting room and a sound-attenuating chamber. NLP and SpandH maintain their own cluster based on SUN GridEngine and Hadoop map-reduce architectures, along with high performance cluster file-systems. Labs for CompBio are provided in the Kroto building housing, e.g., a specialist robot lab for *Marshall*'s evolutionary behaviour lab. VRGraphics run a Reflex Lab that provides support for simulation and 3D visualizations, in addition a new specialist lab for work on haptics and surgical simulation has just opened (our investment of £350k). Additional compute resources are provided by Iceberg, the Uni HPC cluster (~1500 CPU + 3584 GPU cores), and much larger inter-university HPC clusters: the White Rose Grid (+ York & Leeds) and the N8 cluster (with 7 other Northern universities).

e. Collaboration or contribution to the discipline or research base

(i) We have a wide range of **research collaborations:** £18M of funded research over the period has involved collaborating with 77 companies, as well as 190 universities / research institutes. Our exceptional record in EU collaborations - 33 projects since 2008, across all our research groups enables us to collaborate with world-leading groups across Europe. EU projects alone involved 128 universities, 49 companies and 9 charities. National and international collaborations are also supported by RCUK funds, Royal Society, Wellcome Trust etc. For example, large EU projects in NLP, OAK and SpandH have drawn together the best teams in Europe.

In CompBio, extensive collaboration with leading groups in Manchester and Oxford is supported by EPSRC and EU funding. The NST programme grant in SpandH enables us to collaborate closely with Edinburgh and Cambridge – the two other centres of excellence in Speech work. In VT, EPSRC grants with UCL bring together the two largest groups on search-based testing. Outside of Europe we have active collaborations with, e.g., Chicago, Johns Hopkins, NYU, Cornell, Princeton, Queensland.

Research institute collaborations include Rutherford Appleton, MoD, DSTL, EBI Cambridge, EMBL Heidelberg, and **industrial collaborators** include Microsoft, NICTA, Google, Yahoo!, Rolls-Royce, COSTAIN, Quviq, Ericsson, ESL, SAP, BAe Systems, Phillips, IBM, and a leading F1 team.

(ii) Much of our work is **interdisciplinary** by its very nature, and such research lies at the core of our work in CompBio, SpandH, NLP and many of the applications of ML and VRGraphics. Our work on extraction of knowledge from data includes information extraction from biomedical texts, inference of networks from genomic data (NLP, OAK, ML, CompBio). Similarly in the modelling and simulation of natural and computational systems work includes modeling skin cells (CompBio), or complex voice patterns (SpandH), or surgical procedures (VRGraphics) – all of which need a direct dialogue with the other relevant disciplines.

At Sheffield such work is supported by *centres*, *physical co-location* of key groups, and explicit support for new initiatives through *funding* networks. We play a key part in several cross-Faculty centres: **SITraN**: *Sheffield Institute for Translational Neuroscience* (*Lawrence*) founded 2010 as an international centre in neurodegenerative disease. Lawrence has a joint chair with Neuroscience and his research group is embedded in SITraN's new £6M facility. **INSIGNEO**: *Institute for in Silico Medicine* (*Clayton, Harders, Lawrence, Vasilaki, Watton, Romano, Maddock*) established in 2012 through collaboration with Sheffield Teaching Hospitals NHS trust. It focuses on *in silico* medicine with the goal of developing the Virtual Physiological Human as defined by the EU strategic road map. *Harders* and *Watton* were recruited to further our work in this area; since 2012 the institute has already attracted £10M of research funding.

Two new centres will provide further impetus to interdisciplinary work. Together with other departments, we (*Green*, *Moore*) have won some £3M of funding in clinical applications of speech technology. These successes are one driver behind the University's recent £1.3M investment in establishing **CATCH**: *the Centre for Assistive Technology and Connected Healthcare*, based on a Living Lab, with the mission of researching 21st century technologies to enable independent living. The *Sheffield Centre for Robotics* (**SCentRo**) (*Moore, Marshall, Vasilaki, Romano*), supported by a new £1M EPSRC capital grant, provides a focus for robotics research, and will be located in the new Engineering Graduate School from 2014**.** Our work in CompBio also benefits from physical co-location in the **Kroto Research Institute** (*Walker, Marshall, Romano*) that provides state of the art facilities for research at the interface with the Life Sciences.

(iii) We have extensive **collaborations with research users** particularly through EU initiatives that

encourage collaboration with industry (74 industrial partners on current projects). These have led to numerous cases (both strategic and individual) of our work being informed by their concerns and direction. Of course, our industrial funded research projects help us to align our research strategy with theirs directly. Some examples:

In **NLP**, Gaizauskas began researching the potential for comparable corpora to be used in statistical machine translation at the invitation of Tilde, a Latvian language engineering company, with whom he had collaborated previously in two EU projects - the new collaboration has led to two further EU projects and multiple publications (COLING 2012, ACL13). In **OAK**, Rolls-Royce directly funded projects (IPAS Extension, Samulet and Siloet). Spin-out companies (K-Now Ltd and The Floow) have been partners in several projects (X-Media, TRIDS, GrassPortal, etc.) and are exploiting our technologies, and these interactions influence our research direction.

In **CompBio**, research users are clinicians, experimentalists and industry. Active engagement with clinicians includes those at hospitals in Sheffield, Nottingham and London, and experimental biologists at York and Oxford. Industry funding (e.g., Procter and Gamble) and industrial partners in EU projects (Materialise, ANSYS, FluidDA) provide further input to research.

In **ML** Genomatix are partners in two EU projects, we distribute software to research users through Bioconductor packages (2 finished, 1 in preparation). Close collaboration with biologists and clinicians is achieved through embedding part of the ML group in SITraN.

In **VT** Quviq's use of our StateChum tool led to further EU funded work and an EPSRC Kickstart project and use of alternative learning algorithms in our model inference work. In **VRGraphics** application of FLAME-based agent simulation was driven by COSTAIN's need for pedestrian flow models in their design of major transport hubs, and was supported by a RAEng Fellowship.

(iv) Leadership: We were the first UK CS department to form a large speech and hearing group (1986) and one of the earliest to form ML (1998) and comp biology groups (2002), and have one of the largest groups on testing. Staff are involved in international and national meeting organisation to set the agenda and communicate advances to as wide an audience as possible.

Particular *distinctions* include: **Journal editorships of** PLoS one (*Vasilaki, Marshall*), Evolutionary Computation (*Sudholt*), Ass. EIC IEEE Transactions on Pattern Analysis and Machine Intelligence (*Lawrence*), Editor-in-Chief Computer Speech & Language (*Moore*), Ass. Ed IEEE Trans. on Haptics (*Harders*). **Leadership roles in** Chair of the Creative and Digital Industries Sector of the Sheffield City Region Local Enterprise Partnership, and EPSRC Strategic Advisory Panel (*Holcombe*). INTERSPEECH 2009 (*Moore*) **conference chair**; **learned societies:** Vice-chair of IFIP WG 6.1 to 2012 (*Derrick*). **Fellowships** of *Bontcheva* (EPSRC Career Acceleration), *Stevenson, Cohn* (EPSRC advanced), *Romano* (RAEng Industrial), *Hensman* (MRC).

Additional **journal editorships** include: Frontiers in Physiology; JMLR; Journal of Language Resources and Evaluation; Journal of Web Semantics; Ass. Ed of Human Computer Studies; Ed Board for Speech Communication, IJCiNi, Languages. Editor of BioSystems and European Editor for Journal of Visualisation. Ass. Ed ACM Trans. on Applied Perception, Assoc. Ed Virtual Reality and MIT Presence. *Editorial Boards* for IJGCMS, Multimodal User Interfaces, STVR, Extreme and Agile Software Development. *Special issues* of FACS, SCP, IJFCS, BioSystems, Natural Computing, IJUC, Fundamenta Informaticae, Parallel Processing Letters, STVR, IST, SQJ, SPE. **Learned societies:** Chair IEEE Robotics & Automation Committee on Haptics, Vice-President of EuroHaptics Society. Board Director Women's Engineering Society. **Conference chairs**: Mathematical Models in Ecology and Evolution; AISTATS; NIPS Workshops Chair; Founding Editor of JMLR Conference; European Semantic Web Conference; FM'08 Workshop Chair; ABZ/iFM 2012; *Steering committees* of FMOODS; FORTE/PSTV; TESTCOM. Automated Soft Engineering, Lattices and Relations, TAIC PART, SSBSE, TAP, CSTVA, A_MOST, SBST, ICST etc. WorldHaptics; *Area Chair* MICCAI; IEEE Intelligent Robots and Systems; ACL; AISB, in addition to membership of many **Programme Cttees**. **Invited keynotes**: TEMTIA; Bioengineering; SSSW12; Asian Semantic web conference; BIS; RelMiCS; RAMICS; ALCOP; WIN VR; Virtual Reality in Medicine; Mechatronics and IT, etc. Finally, **best paper awards** at GECCO; PPSN; ISWC and more.

We are well placed to sustain our leadership, with our new staff, including 7 ECRs already making strong contributions via fellowships, journal editorships, conference chairs, programme committees, and keynotes.