



# Research Newsletter

## School of Computer Science



### In this issue

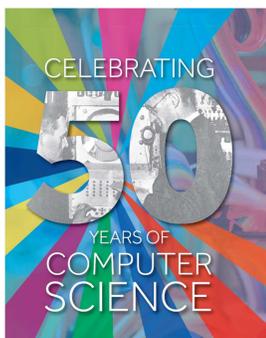
- p.2 Head of School Editorial and Research News
- p.4 Feature: #BritainBreathing
- p.5 Spotlight: Publication Success
- p.8 Grants and Awards

### Contact us

**School of Computer Science**  
**The University of Manchester**  
**Kilburn Building**  
**Manchester**  
**M13 9PL**  
**UK**

Email: [researchsupportcsm@manchester.ac.uk](mailto:researchsupportcsm@manchester.ac.uk)

Web: [www.cs.manchester.ac.uk](http://www.cs.manchester.ac.uk)



### About us

At The University of Manchester, we have one of the longest established schools of computer science in the UK and one of the largest. We are constantly building on our strong research history with research groups operating across the spectrum of computer science, from fundamental theory and innovative technology, through novel hardware and software systems design, to leading-edge applications.

The School is consistently ranked highly; top 5% in the UK (REF2014, GPA); assessed as *the best environment* in the UK for computer science and informatics research (REF2014); 7th in the UK by ARWU 2014 and the expertise and achievements of our staff are well-recognised internationally.

# Editorial by the Head of School

We've had a very successful year in the School of Computer Science, with some excellent publications, continued success in gaining funding at a total value of £3.8M, and continued success in getting EPSRC Impact Acceleration Account awards, which help our research make an impact. We are looking forward to participation in the CityVerve project, Manchester being the European City of Science and expanding our academic staff by more new appointments in 2017.

I'd also like to take this opportunity to thank our research staff and support staff, who make all of this possible. Take a look below at some of the research and discoveries that are happening at the moment.



Professor Robert Stevens

## News

### Professor Sophia Ananiadou interviewed by Pharma Technology Focus.

The Director of the National Centre for Text Mining discusses the huge benefits their work could have to clinical research.

*"The key is that, because software tools rather than people are doing the analysis, the amount of time, and therefore money, which can be saved by researchers is phenomenal."*

*"Text mining doesn't necessarily give you all of the answers, but its potential to speed up the research process and suggest more promising paths for drug discovery is huge."*

Full article: <http://www.nridigital.com/pharma-technology-magazine/june-2016.html?vw=s%2FPharma%2520Technology%2520Focus%2F8e39fe4d-1815-5813-871a-d40dd3c7b8a6%2FPTF%25201606%2Ftextmining.html>

### NaCTeM puts in a strong showing at SemEval 2016.

The 10<sup>th</sup> International Workshop on Semantic Evaluation – SemEval 2016, was held in San Diego, California on the 16<sup>th</sup> and 17<sup>th</sup> June. The event was co-located with the 15th Annual Conference of the North American Chapter of the Association for Computational Linguists.

SemEval is a series of evaluations of natural language processing systems, exploring the nature of meaning by comparing different algorithms intended to mimic human language understanding. This year, NaCTeM participated in the Semantic Textual Similarity task and was ranked 4th out of 43 teams. The competing systems were given over 9000 pairs of sentences and were asked to return a value indicating a level of equivalence of their meaning. Performance was assessed by computing the correlation with scores assigned by human judges and NaCTeM reached a very high level of agreement (0.7486).

A description of the approach can be found in the following paper:

P. Przybyła, N. T. H. Nguyen, M. Shardlow, G. Kontonatsios and S. Ananiadou, "NaCTeM at SemEval-2016 Task 1: Inferring sentence-level semantic similarity from an ensemble of complementary lexical and sentence-level features," *Proceedings of the 10th International Workshop on Semantic Evaluation (SemEval 2016), San Diego, USA, 2016. PURE ID: 34724197*



### Welcome to Professor Junichi Tsuji.

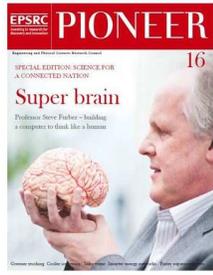
The School is delighted to Welcome Prof Junichi Tsujii who is a newly appointed part-time professor in Natural Language Processing. Prof Tsujii is also director of the Artificial Intelligence Research Center at the National Institute of Advanced Industrial Science and Technology, Japan (AIST), and aims to build links between the University and AIST across a range of topics.

Prof. Tsuji, also a Scientific Advisor to NaCTeM, has recently been awarded the Okawa prize, for his pioneering contributions to research in computational linguistics and natural language processing including machine translation and text mining. The Okawa Prize is awarded to pay tribute to and publically recognise outstanding, international level contributions to research, technological development and business in the information and telecommunications fields.

<http://www.okawa-foundation.or.jp/en/activities/prize/list.html>



## Professor Steve Furber is featured on the cover of EPSRC Pioneer Magazine #16.



The ICL Professor of Computer Engineering is interviewed in depth by EPSRC's quarterly research review, discussing the development of the ARM microprocessor, SpiNNaker and the Human Brain Project.

*"The new generation of embedded systems will change the way we work, do business, shop, travel, and care for ourselves through a wide range of applications where low energy consumption and reliability are central. "Ultimately, these systems will shape the emergence of a new digital society for the 21st century."*

<https://www.epsrc.ac.uk/newsevents/pubs/pioneer16/>

## Daniel Dresner in the news.



Expert in Cyber Security, Dr. Daniel Dresner has appeared in the media discussing some of the most high-profile computer hacks of recent times.

In April he was interviewed by the BBC World Service Business Matters<sup>1</sup> Programme discussing the Bangladesh cyber bank raid, the largest such computer hack in history.

Earlier that month he was asked by Wired.co.uk<sup>2</sup> to comment on the flawed security protocols at Panama based law firm Mossack Fonseca that allowed the leak of the so-called Panama Papers.

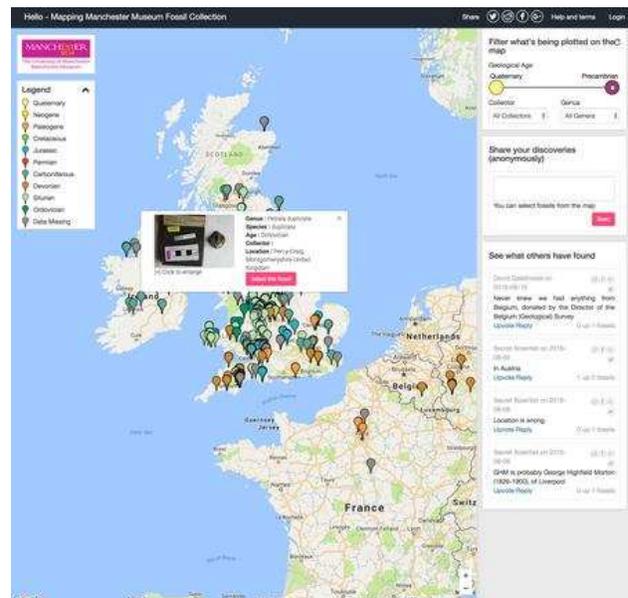
## Making discoveries on a fossil collection through citizen science

As part of his MSc project at the School of Computer Science, Olivier Staub has developed a citizen science application that displays on a map the fossil collection of the Manchester Museum. The goal of the application is to allow the members of the community to submit the discoveries made through the exploration of the map. A discovery doesn't have to necessarily be ground-breaking but can be any regularity or irregularity found such as "in the late 50s, John Smith collected trilobites in Wales".

See the site at <https://natureslibrary.co.uk/map>

*"Our goal as Human-Computer Interaction specialists is to understand how people make discoveries while they explore maps. Ultimately, this understanding will enable the development of tools that facilitate discoveries using geographical information systems. We are looking for participants so your contributions are key to the success of the project - remember you don't have to be an expert to take part!"*

This is a collaborative project between the School of Computer Science and the Manchester Museum: Olivier Staub is supervised by Markel Vigo and he's working under the advice of David Gelsthorpe, the Curator of Earth Science Collections.



# Feature: Britain Breathing



**#BritainBreathing is an exciting new citizen science project that aims to engage the UK public to act as 'citizen sensors' to help scientists discover more about seasonal allergies such as hay fever or asthma.**

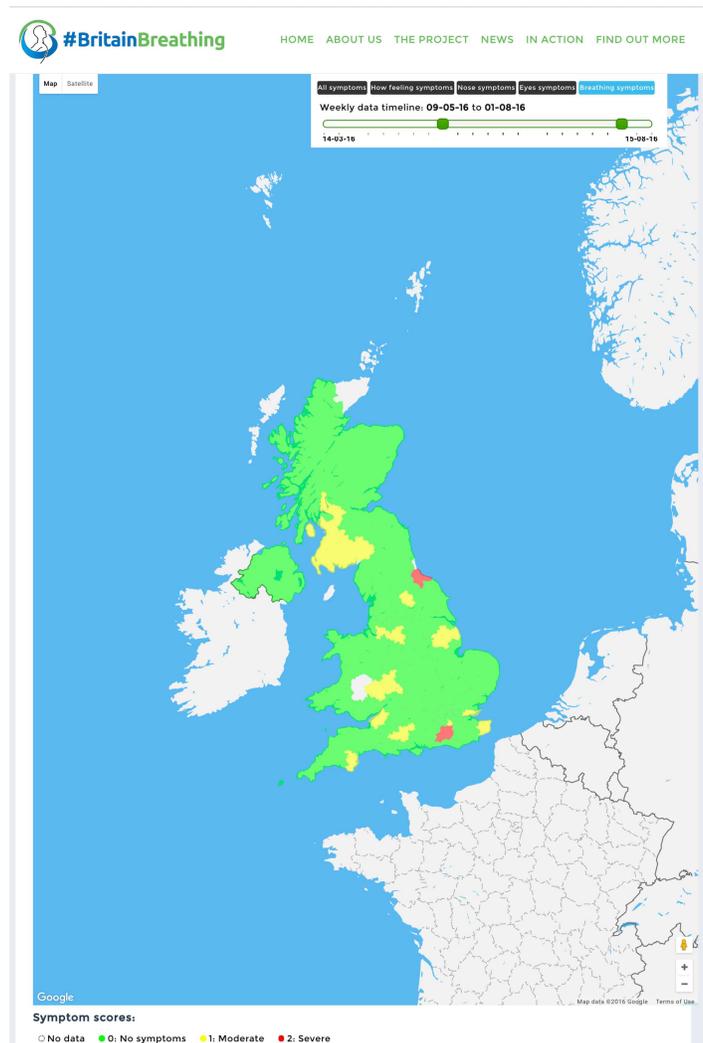
This is a joint project between the British Society for Immunology, Royal Society of Biology and The University of Manchester – including Dr. Caroline Jay, Prof. Andy Brass and Dr. Markel Vigo from the School of Computer Science. The aim of this project is to better understand (both on a national scale and for the individual) when allergy symptoms are occurring and what the triggers for these might be.

The team have developed a free-to-use app that will allow the public to record their allergy symptoms in a simple and straightforward way and then safely share that data with the project team. This large data set (capturing information on timing and location of allergy symptoms) can then be combined with other publicly available data (such as weather, pollen or pollution statistics) to build a better understanding at a national and personal level of allergy and allergy triggers. From these data, we can build a clearer picture of the pattern and frequency of allergy incidence across the UK.

Markel Vigo represented the project at the recent EuroScience Open Forum, held in Manchester: *“As part of ESOE, the #BritainBreathing project presented an allergy symptom visualisation widget at the Exchange Court in Manchester Arndale. The widget displays the symptoms reported by the citizen scientist who are using the #BritainBreathing app and allows to explore the map of the UK filtering the data per symptom, week and region. The application was well received and has become the public face of the #BritainBreathing project, which has collected so far 15K observations since April.”*

The widget is available

at: <http://britainbreathing.org/index.php/data-visualisation/>

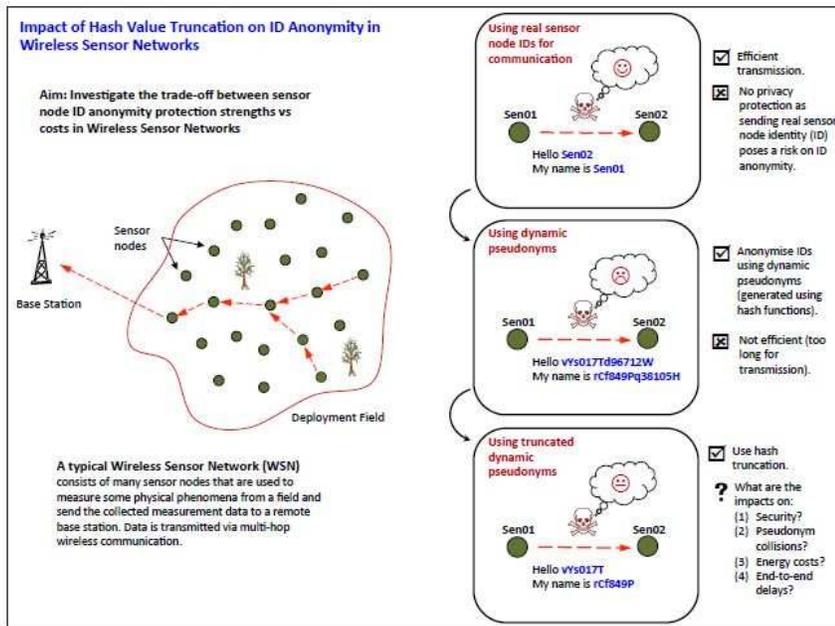




**'Impact of Hash Value Truncation on ID Anonymity in Wireless Sensor Networks.'** Ahmed Al-Riyami, Ning Zhang, John Keane, *Ad HOC NETWORKS*; 45; PP 80–103<sup>5</sup>. PURE ID: 22516228

"This paper has reported a comprehensive study of the implications of using hash values and hash value truncations in preserving node ID (identity) privacy on security and performance in a wireless sensor network context. It has

considered a comprehensive list of factors including easiness in launching security attacks, hash value collisions, the implication of the collisions on the computational costs, the trade-offs in terms of energy costs and end-to-end packet delivery delays. Based on two existing ID anonymity schemes, the paper uses both theoretical analysis and simulation study to provide some interesting findings and insights on the trade-offs between achieving privacy and performance implications. These findings are of great interests not just to sensor network community, but also emerging fields such as the Internet of Things (IoT), as hash functions are also used to achieve other security properties and in many other fields including IoT. "

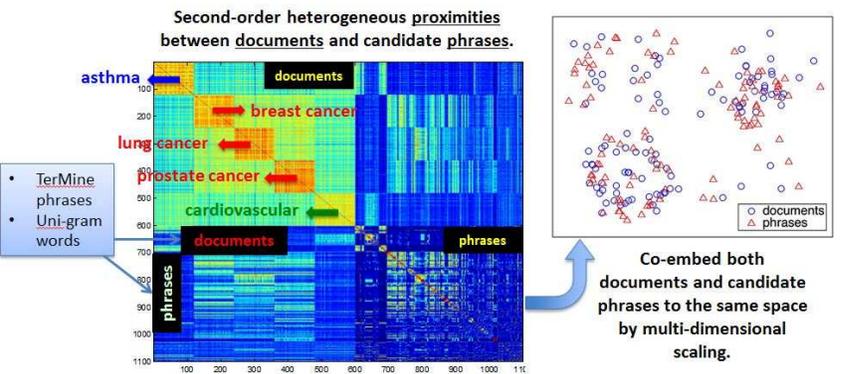
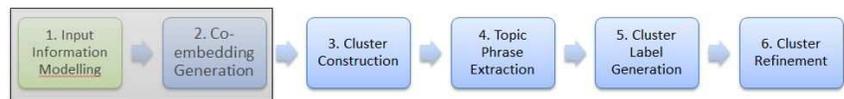


**'Descriptive Clustering via Discriminant Learning in a Co-embedded Space of Multi-level Similarities.'** Mu, T., Goulermas, J., Korkontzelos, I., and Ananiadou, S., *Journal of the Association for Information Science and Technology*<sup>6</sup>, 67(1), 106-133.

This paper suggests a novel method for descriptive document clustering, called CEDL, which discovers and groups automatically, semantically interrelated documents and at the same time summarises their content with meaningful cluster labels. CEDL uses several processing stages: co-embedding, clustering, classification and ranking techniques to group the documents and choose the best and more informative labels.

which deal with big textual data. Clustering is used in text mining to discover diverse groups of semantically related documents and making different interrelated informational views in big textual data more accessible, searchable and useful by both humans and machines. Our research discovers semantic topics in a complex and heterogeneous problem space. Thus, it addresses challenges of fragmentation since knowledge is distributed and hidden in big textual sources.

In a novel way, it analyses jointly documents and phrases (concepts automatically identified from documents) to improve the quality of document clusters and to find labels that best describe a document cluster. Comparisons with popular descriptive clustering algorithms (descriptive k-means and concept-driven clustering based on matrix factorisation) show that CEDL generates more informative labels and higher quality document clusters. The method has been used as part of a text analysis framework to support search. It has been applied to different domains such as clinical trials, newswire and public health systematic reviews



**'Large-scale simulations of plastic neural networks on neuromorphic hardware.'** Knight, James Courtney and Tully, Philip Joseph and Kaplan, Bernhard A and Lansner, Anders and Furber, Steve B, *Frontiers in Neuroanatomy*, Vol.10, No. 37, 2016<sup>7</sup>. PURE ID: 24034365

This paper describes the results of an international collaboration to test an advanced concept for learning in a neural network - modelled on a theory of how the brain might learn things - on a computer system that is itself designed on brain-inspired principles. The model incorporates 51 million plastic synapses - these are the connections between the neurons where learning takes place - and is believed to be the largest plastic neural network ever to have been run on - neuromorphic - that is, brain-inspired - hardware.

The paper's lead author - Jamie Knight - is a PhD student in his final year at the University of Manchester, and he has carried out the detailed modelling work described in the paper as well as coordinating the international collaboration with senior colleagues at KTH Stockholm and the Zuse Institute, Berlin.

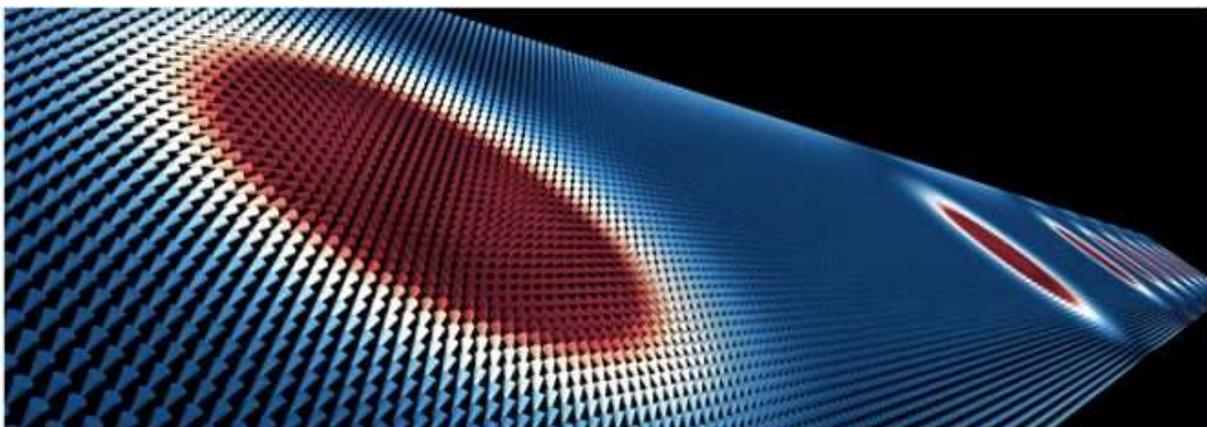


**'Additive interfacial chiral interaction in multilayers for stabilization of small individual skyrmions at room temperature. Nature Nanotechnology.'** Moutafis, C., Moreau-Luchaire, C., Moutafis, C., Reyren, N., Sampaio, J., Vaz, C. A. F., Horne, N. V., Bouzehouane, K., Garcia, K., Deranlot, C., Warnicke, P., Wohlhüter, P., George, J., Weigand, M., Raabe, J., Cros, V., Fert, A., *Nature Nanotechnology* 11, 444-448<sup>8</sup>, 2016 PURE ID: 27858647

This paper describes the first observations of nanoscale chiral magnetic skyrmions at room temperature. It is an experimental breakthrough in the burgeoning field of spintronics, or spin-based electronics, which studies the role of the electron spin in solid state physics; it also studies the technological applications that exploit the electron's spin properties, as well as the electron's charge used in conventional electronics.

Spintronics-based devices can enable non-volatile power efficient functions, like in MRAMs (Magnetic Random Access Memories), as they do not need an electric current to retain their "spin" or magnetic state. Nanoscale spintronics devices could in principle be smaller, faster, more robust and energy efficient than current technologies for data storage. Skyrmions resemble small magnetic vortices and can occur in many materials. In this work, tiny skyrmions, in the nanoscale regime, were demonstrated in a technologically relevant material, at room temperature. In order to achieve this, the

research team tailor-designed the necessary functionality in thin film materials and used X-ray imaging to determine the presence of chiral magnetic skyrmions. In today's magnetic hard drives, data are stored in magnetic domains (with magnetic spins aligned in the same direction), but there are fundamental limits to how small such domains can become and how tightly they can be packed together. Skyrmions however could be made much smaller and could therefore be used in data storage devices with much higher density. An additional key advantage is that due to the skyrmions' smaller size and robustness against perturbations and defects, they can be packed efficiently and moved rapidly along nanowires in so-called racetrack future MRAM type of memories. These memories would be more robust as they would not require moving mechanical parts seen in today's hard drives, would retain information when the power is switched off and be more energy efficient, which is paramount in our increasingly technological Information Society.

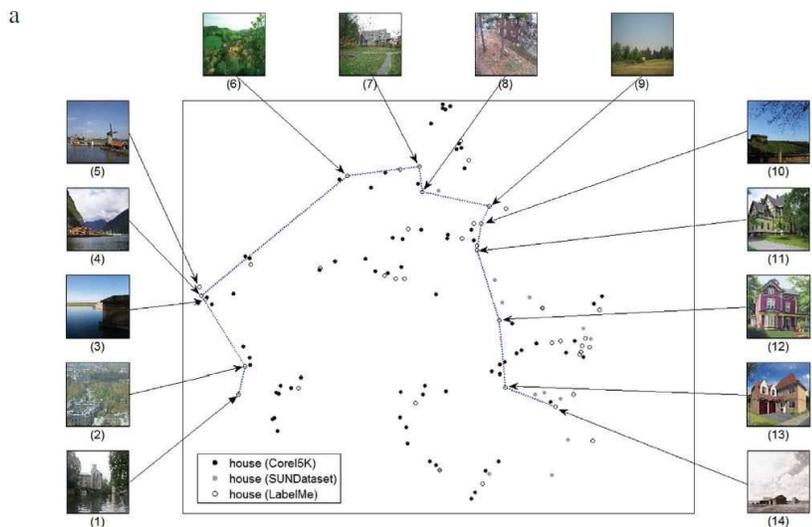


Multimedia information retrieval and understanding demand modelling contextualized semantics underlying the multimedia description. Descriptive terms, e.g. labels and tags, are often used to annotate multimedia. Unlike natural

languages, such terms are neither limited to words nor have syntactic orders when they are jointly used to describe images, music track and video clips. How to model the concept formulated by a term in context of its co-occurring

terms remained unclear until our work has been done. In this paper, we innovatively formulate modelling contextualised semantics underlying a term as an unsupervised learning problem that embeds all concepts defined by terms along with their context in a semantic representation space where a term has multiple polysemous representations when it co-occurs with different terms and the distance between embedded concepts reflects their semantic similarity (as exemplified in the figure).

Furthermore, we propose a novel deep neural architecture to carry out this learning task. The resultant semantic representation allows for not only modelling contextualised semantics of known terms in a given vocabulary during learning but also inferring that of unknown or out-of-vocabulary terms from co-occurring terms without re-learning. The proposed approach has been thoroughly evaluated in different domains and leads to non-trivial applications.



(1) bridge, <b>house</b> , tower	(8) tree, grass, <b>house</b> , window, bench
(2) city, bridge, river, <b>house</b> , castle, park, tower	(9) sky, tree, grass, fence, <b>house</b>
(3) sky, water, <b>house</b>	(10) sky, tree, <b>house</b>
(4) mountain, sky, water, <b>house</b>	(11) sky, tree, grass, <b>house</b> , road
(5) sky, water, <b>house</b>	(12) city, sky, tree, <b>house</b> , path, window, door, garden, lawn
(6) tree, forest, field, <b>house</b>	(13) sky, grass, <b>house</b> , window, street, door
(7) tree, grass, <b>house</b>	(14) sky, land, ground, <b>house</b>

## Grants and awards

The School of Computer Science has been awarded over **£13 million** in external funding for research over the last two years. Much of the research involves working in collaboration with others across the University and all over the world.

Here are just some examples of recent research funding awarded in the School.

### LAMBDA: Learning Algorithms for Modularity in Broad and Deep Architectures

Dr. Gavin Brown and Prof. Mikel Lujan

Funding body: EPSRC

Award amount: **£521,154**

Do you know how Facebook recognises faces in images? Do you know how your iPhone understands your speech? The secret behind each of these is a technology called "Deep Learning", which uses biologically-inspired algorithms called "neural networks".

Over the next decade, society will become more reliant on this technology. But... these algorithms require an IMMENSE amount of computing power, and therefore electricity, and for example can take many weeks to learn a given task.

The LAMBDA project explores an approach to deep learning which is not just deep, but also broad - hence "Learning Algorithms for Broad and Deep Architectures". We aim to (1) make "broad" models that are faster/easier to learn, and as a consequence (2) reduce the energy consumption. Our approach builds upon previous award winning research by the PI, in exactly this area. If successful, we will be able to reproduce the same abilities as current deep neural networks, but with a significantly reduced energy consumption, and whilst learning such architectures a significantly easier task for scientists.

## **IAA C&F with MDSAS: Evidence-based Decision Support for Haemophilia Treatment**

**Prof. John Keane**

**Funding body: Medical Data Solutions and Services Ltd (MDSAS) & EPSRC - IAA Concept and Feasibility Study**

**Award amount: £58,059**

*Haemtrack* is a highly successful online national system for Haemophilia patients who, due to treatment advances, largely treat themselves outside hospital.

*Haemtrack*, developed by MDSAS, allows patients to record their therapy; if a patient uses treatments in response to an injury or spontaneous bleed, *Haemtrack* will collect detailed information specific to this. Healthcare professionals can view this information in real-time and use it to monitor the health of patients outside hospital and also to review with patients during clinics.



The aim of this project is to assess the feasibility of developing and incorporating advanced and innovative data analysis and decision support techniques, termed here Evidence-Based Decision Support (EB-DS), into *Haemtrack*. The EB-DS technology will monitor patient treatments and provide personalised dosing recommendations to healthcare professionals that monitor those patients.

## **DISRUPT: Factories of the Future**

**Dr. Rizos Sakellariou** (led by CENTRO RICERCHE FIAT SCPA , Italy)

**Funding Body: H2020 Research and Innovation Action**

**Award amount: £231,671**

Industry 4.0 or the fourth industrial revolution is the next developmental stage in the organisation and management of industrial processes and the entire manufacturing value chain. Traditional manufacturing and production methods are in the throes of a digital transformation. By blending the real and the virtual production world, it is now possible to connect all parts of the production process: machines, products, systems, and people. In that regard, ICT-based systems and service platforms will play a major role in this transformation; they bring connectivity and enable monitoring, analysing, simulating, optimising and controlling production entities and processes, thus creating a virtual copy of the physical world and facilitating decentralised structures through Cyber-Physical Systems (CPS). Over the Internet-of-Things (IoT), CPSs communicate and cooperate with each other and humans in real-time. Via the Internet-of-Services (IoS), both internal and cross-organisational services are offered and utilised by participants of the value chain.

DISRUPT is a consortium led project, which aims to spearhead the transition to the next-generation manufacturing paradigm by facilitating the vision and execution of a "Smart Factory". In the traditional rigid factory, mass-production is the goal and cost-effectiveness is the drive; this has led many manufacturers to relocate to lower cost regions, and others to totally outsource

their production. The new era of manufacturing asks for optimised plants and manufacturing chain networks, transforming them into profitable innovation centers. It requires flexible factories that can be quickly "reprogrammed" to provide faster time-to-market responding to global consumer demand, effectively addressing mass-customisation needs and bringing life to innovative new products.

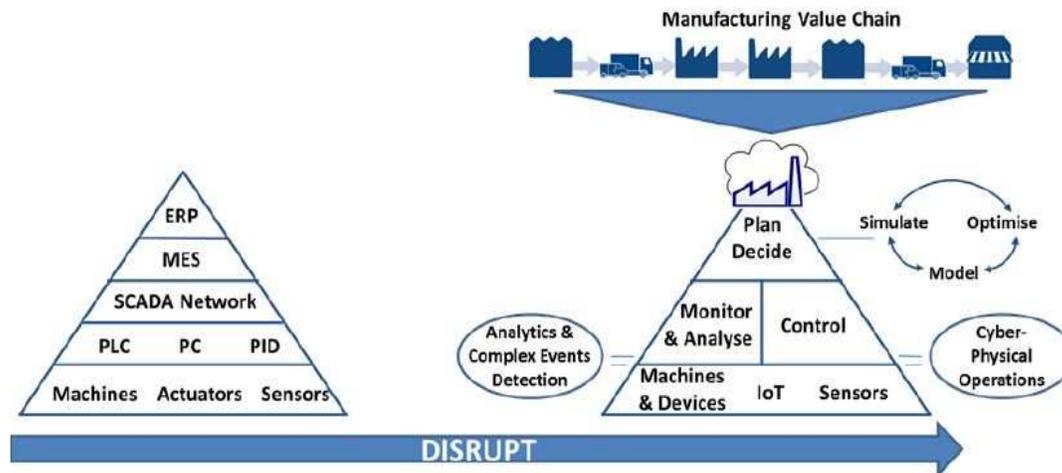


Figure 1 The DISRUPT Concept: Disrupting the traditional automation pyramid (Industry 4.0)

# Research Programmes in Computer Science

By choosing to study at Manchester you will be joining one of the most innovative and successful schools of computer science in the world. Manchester saw the birth of computer science, with the creation of the world's first stored-program computer. We continue to work on pioneering research including cutting-edge image processing software, new technologies to exploit the power of the web, medical imaging software and low-power chip design. The School incorporates ten research groups operating across the spectrum, from fundamental theory and innovative technology, through novel hardware and software systems design, to leading edge applications.

Although a research degree is traditionally the route to an academic post, and is often an essential qualification for a research career, it is now frequently sought by industry. Research graduates are a step ahead in the race for jobs: the communication and analysis skills integral to a research degree make them highly employable individuals.

For more information, please contact [pgr-admissions@cs.man.ac.uk](mailto:pgr-admissions@cs.man.ac.uk) or visit [www.cs.manchester.ac.uk/study/postgraduate-research](http://www.cs.manchester.ac.uk/study/postgraduate-research)

## Links

1. Daniel Dresner on BBC World Service:  
[http://www.bbc.co.uk/programmes/p03rh5lt?gator\\_td=ru2tvkkgk9AfXwzjmeAZsgDZLY5Ycdv3M0o7oaPJH6gdhZzPPvDS1Kyr7a3at6Pja2W2WgldO5N5lpcDZTeEJw10LunxW0Bqur5eCTiUVQU1lhgRIZMefVmITRDBxg%2fNZ1ftllQ207x359%2fhf071pFBnYpbzfnmF8M5qunmw8R4SWF%2fxhE%2b9IQI2KbF7VBNI#playt=0h1m49s](http://www.bbc.co.uk/programmes/p03rh5lt?gator_td=ru2tvkkgk9AfXwzjmeAZsgDZLY5Ycdv3M0o7oaPJH6gdhZzPPvDS1Kyr7a3at6Pja2W2WgldO5N5lpcDZTeEJw10LunxW0Bqur5eCTiUVQU1lhgRIZMefVmITRDBxg%2fNZ1ftllQ207x359%2fhf071pFBnYpbzfnmF8M5qunmw8R4SWF%2fxhE%2b9IQI2KbF7VBNI#playt=0h1m49s)
2. Daniel Dresner in Wired magazine:  
<http://www.wired.co.uk/article/panama-papers-mossack-fonseca-website-security-problems>
3. Bias in the reporting of sex and age in biomedical research on mouse models:  
<https://elifesciences.org/content/5/e13615>
4. The Landing Gear System in Multi-Machine Hybrid Event-B:  
<http://link.springer.com/article/10.1007%2Fs10009-015-0409-7>
5. Impact of Hash Value Truncation on ID Anonymity in Wireless Sensor Networks:  
[https://pure.manchester.ac.uk/admin/editor/dk/atira/pure/api/shared/model/base\\_uk/researchoutput/editor/contributiontojournaleditor.xhtml](https://pure.manchester.ac.uk/admin/editor/dk/atira/pure/api/shared/model/base_uk/researchoutput/editor/contributiontojournaleditor.xhtml)
6. Descriptive document clustering via discriminant learning in a co-embedded space of multilevel similarities:  
<http://onlinelibrary.wiley.com/doi/10.1002/asi.23374/abstract>
7. Large-scale simulations of plastic neural networks on neuromorphic hardware:  
<http://journal.frontiersin.org/article/10.3389/fnana.2016.00037/full>
8. Additive interfacial chiral interaction in multilayers for stabilization of small individual skyrmions at room temperature. Nature Nanotechnology:  
<http://www.nature.com/nnano/journal/v11/n5/full/nnano.2015.313.html>
9. Learning contextualized semantics from co-occurring terms via a Siamese architecture:  
<http://www.sciencedirect.com/science/article/pii/S0893608016000058>

For past issues of the School of Computer Science Research Newsletter see [www.cs.manchester.ac.uk/our-research/news/](http://www.cs.manchester.ac.uk/our-research/news/)