

Report on YERUN Research Workshop

Big Data, Data Analytics and the Digital Economy

January 25 – 26, 2018
University of Essex



Organising Committee:

Professor Corrado Cerruti, University of Rome Tor Vergata, Italy
Professor Javier Prieto, Universidad Carlos III de Madrid, Spain
Professor Christine Raines, University of Essex, UK

CONTENTS

Introduction	3
Programme	4
Summary of workshop.....	7
Outcomes.....	9
Participant biographies.....	12

INTRODUCTION

This first YERUN research workshop was the inaugural event in a series of workshops to exchange research ideas, gain a better understanding of complementary expertise in the thematic priorities and develop joint research collaborations between YERUN partners.

This workshop focused on the YERUN priority area of **Big Data, Data Analytics and the Digital Economy** which draws together expertise in analytics, data storage, managing and curating data, creative innovative data sets, developing new scientific techniques to analyse data and generate knowledge, analysing the impact on business models and business processes, educating the next generation of data scientists, spreading basic knowledge on data management among students of different fields, education of the next generation of data scientists and applying scientific knowledge to the real world through understanding of human behaviour.

There were two broad aims of the workshop, to enable participants to:

- Exchange research ideas and gain a better understanding of complementary expertise in the thematic priorities identified within the network, and
- Develop ideas and for new research projects and collaborations and in particular action plans to take forward these ideas for the H2020 funding call

The workshop brought together 30 academics from 15 YERUN partners with a broad range of expertise from computer science to cultural heritage. Over the two days of the workshop participants were asked to undertake concentrated discussions to make connections and develop concrete plans for research projects with the aim of developing successful grant applications. An external facilitator guided the discussions and ensured the event kept to time and purpose.

This report brings together information about the event, its participants and the outcomes of the workshop.

PROGRAMME

YERUN Research Workshop

Big Data, Data Analytics and the Digital Economy

Facilitator: Carol Sherriff

TIMETABLE

Day 1: Thursday January 25th

Timing	Activity
2.30pm	OPTIONAL: Campus tour (meet outside Wivenhoe House Hotel)
3.30pm	Welcome refreshments
4.00pm	<p>‘Getting to know you’ session</p> <p>Welcome from workshop hosts: <i>Professor Corrado Cerruti, University of Rome Tor Vergata</i> <i>Professor Javier Prieto, Universidad Carlos III de Madrid</i> <i>Professor Christine Raines, University of Essex</i></p> <p>Formal workshop opening by YERUN President: <i>Professor Juan Romo, Universidad Carlos III de Madrid</i></p> <p>Agreement on purpose and programme for the two days</p> <p>Brief introduction to the work of the 15 YERUN partners (4 minutes each)</p>
5.10pm-6.00pm	<p>Sharing preliminary ideas</p> <p>Whole group and small group discussions to identify key themes and ideas to be taken forward on day 2.</p> <p>Prioritising of ideas to be developed on day 2 – criteria:</p> <ul style="list-style-type: none"> • Each idea needs a sponsor/leader who will convene the discussion on day 2 • The involvement of more than one partner and • Preferably connection to one of the H2020 themes
7.00pm	Dinner in Signatures private dining room, Wivenhoe House Hotel

Day 2: Friday January 26th

Timing	Activity
8.00am-9.00am	<p>OPTIONAL parallel sessions:</p> <p>Overview of Horizon 2020 funding - <i>Juliet Craig, EU Research Development Manager, University of Essex</i></p> <p>OR</p> <p>Campus tour (meet outside Wivenhoe House Hotel)</p>
9.00am	Refreshments
9.15am	Welcome from <i>Professor Anthony Forster, Vice-Chancellor, University of Essex</i>
9.30am	<p>Review of key research themes</p> <ul style="list-style-type: none"> • Visual display of key research themes on flipcharts • 1 minute introduction of each theme by group leader • Identification of any gaps • Thematic groups form around each research idea
11.00am	Break with refreshments
11.30am	<p>Research theme discussion groups (1)</p> <p>Research group discussions to develop ideas, collaborations and how the ideas might be developed</p>
1.00pm	Lunch
2.00pm	<p>Research theme discussion groups (2)</p> <p>Research discussion groups work further on their action plans to include:</p> <ul style="list-style-type: none"> • Overall objectives • Key actions and activities • List of responsibilities • Key dates
3.20pm	Break with refreshments

3.30pm	Presentation of commitments <ul style="list-style-type: none">• Each group makes a brief presentation of their commitments/action plan• Opportunity to circulate and comment on different groups work• Return to discussion groups to review comments and agree on next steps
4.00pm	Conclusion and thanks

SUMMARY OF WORKSHOP

Preparatory work

To ensure the workshop was focused on the development of joint projects, in the weeks prior to the workshop a list of relevant Horizon 2020 grant opportunities was circulated to participants. Everyone was asked to provide their initial ideas for projects and collaborations. These were collated, along with biographies and photographs, and provided in delegate packs at the workshop.

Day One

The workshop kicked off at 3.30pm with many participants arriving straight from a tour of a chilly University of Essex campus to be warmed up with tea and coffee and some intensive networking. The event was formally opened shortly before 4pm with a welcome from Professor Juan Romo, Rector of Universidad Carlos III de Madrid and YERUN President who noted the historic nature of this first YERUN Research Workshop.

Following welcomes from the three workshop convenors - Professor Corrado Cerruti, University of Rome Tor Vergata; Professor Javier Prieto, Universidad Carlos III de Madrid; Professor Christine Raines, University of Essex – the participants were immediately put to work.

With just 8 hours of scheduled working time during the workshop, the challenge was to complete introductions of all participants and their research interests within an hour. To facilitate this each participating institution had been asked to provide a single presentation – to last no more than 4 minutes – which introduced their colleagues and their expertise. Supported by some strict timekeeping this high speed format was adhered to and after an hour, the group had identified some common themes: education, health, business & finance, social science and humanities, and smart cities. Participants formed into groups to discuss the key issues and after just 30 minutes of discussion, were asked to report on the common ground. The feedback reflected a range of overlapping interests:

Education: the group represented a number of allied concepts broadly split across three domains: museums, arts and heritage; social media, and; universities. One specific project was mentioned for the latter domain: how to use data to help students select and complete university courses, and especially how to identify at an early stage those students who are struggling, in order to help them and reduce attrition rates.

Health: key interests in this interdisciplinary group included: solutions for wellbeing in cities; using clinical health data for prediction, prevention and personalised medicine, and; use of data around genomics and other “omics”.

Business and finance: initial ideas were focused on the use of more and better data to improve decision making, however they were also concerned about how to include all expertise into possible projects.

Social science and humanities: this group highlighted the use of data and AI to improve public services and make better decisions, but emphasised the challenges of ensuring the decision-making process could be documented and understood, and fully accountable. The use of textual data analysis was also suggested as a possible topic, but it was not clear how this could be effectively implemented and used.

Smart cities: this group focused on improving the quality of life of people in cities by using data from a wide range of sources to understand, for example, patterns of energy consumption, mobility patterns and barriers to mobility.

In addition two cross-cutting themes – relevant to all topics - were identified: cybersecurity and ethics, and; analytics. With this initial work complete, the group were headed to dinner to continue their discussions informally.

Day Two

The day started with more tea and coffee, and a welcome from Professor Anthony Forster, the Vice-Chancellor of the University of Essex.

Following this, the focus was on transitioning from joint areas of interest to joint areas of collaboration on specific H2020 research projects. After a concentrated 20 minute review of the calls, participants were asked to indicate their interest in each of the forthcoming calls. From this exercise three broad areas of interest groups identified: health; public services, and; energy. Three thematic working groups were formed to discuss the calls and potential responses to them and the work continued until lunch time.

As the three groups were all at various stages of project planning, after the lunch break a brief discussion was held on additional themes which participants might like to explore. This reflected on some of the topics discussed the previous day: ethics and privacy; education; wellbeing, urbanisation and environment, and; analysis and methods. As a result during the afternoon session the groups were more fluid, with some participants forming smaller working groups on these topics and some of the morning groups continuing their discussions.

The hard work continued until 3.30pm – just 24 hours after the start – with all groups reporting back on their exchanges. Each group was asked to identify clear actions, project convenors, and potentially a time frame for future actions. This feedback, reviewed and refined by email after the event, is captured in the next section *Outcomes*.

OUTCOMES

Project proposals

Call	People involved
HEALTH	
DS-TDS-01-2019: Smart and healthy living at home (IA) Deadline: 14 November 2018	YERUN Health Group aims to be meet/skype/phone conference by end of February to discuss further UAM: Victor Rubio, Estrella Pulido, Doroteo Torre Toledano Ulm: Jan Beyersmann Essex: Berthold Lausen, Dimitri Ognibene CarlosIII: Anibal Figueiras-Vidal Antwerp: Bart Goethals
SC1-BHC-13-2019: Mining big data for early detection of infectious disease thread driven by climate change and other factors (RIA) Deadline: 16 April 2019	Essex: Berthold Lausen, Dimitri Ognibene Ulm: Jan Beyersmann Nova: Pedro Barahona UPF: Miguel A Mayer
SC1-DTH-01-2019: Big Data and Artificial Intelligence for monitoring heath status and quality of life after cancer treatment (RIA) Deadline: 24 April 2019	Essex: Berthold Lausen, Dimitri Ognibene Ulm: Jan Beyersmann UPF: Miguel A Mayer UAM: Victor Rubio, Estrella Pulido, Doroteo Torre Toledano Maastricht: Evgueni Smirnov CarlosIII: Anibal Figueiras-Vidal
PUBLIC SERVICES	
DT-TRANSFORMATIONS-02—2019: : Transformative impact of disruptive technologies in public services (RIA) Deadline: 14 March 2019	Essex: Maria Fasli mfasli@essex.ac.uk UAM: Ignacio Criado ignacio.criado@uam.es DCU: Markus Helfert markus.helfert@dcu.ie BREMEN: Sebastian Haunss haunss@uni-bremen.de UAB: Daniel Franco daniel.franco@uab.cat UAM: Victor Rubio victor.rubio@uam.es ESSEX: Paola Di Giuseppantonio Di Franco pd17425@essex.ac.uk NOVA: Roberto Henriques roberto@novaims.unl.pt BRUNEL: Ashley Braganza ashley.braganza@brunel.ac.uk SDU: Peter Schneider- Kamp petersk@imada.sdu.dk

ENERGY	
<p>DT-ICT-11-2019: Big Data solutions for Energy (IA)</p> <p>Deadline: 2 April 2019</p>	<p>NOVA: Mauro Castelli mcastelli@novaims.unl.pt SDU: Mikkel Baun Kjærgaard mbkj@mmmi.sdu.dk UC3M: Francisco Javier franciscojavier.prieto@uc3m.es UTV: Simone Borra borra@economia.uniroma2.it, Valeria Cardellini cardellini@ing.uniroma2.it, Corrado Cerruti corrado.cerruti@uniroma2.it</p>
<p>LC-SC3-EE-13-2018: Enabling next-generation of smart energy services valorising energy efficiency and flexibility at demand-side as energy resource (CSA)</p> <p>Deadline: 4 September 2018</p>	<p>UTV (Coordinator): Simone Borra borra@economia.uniroma2.it, Valeria Cardellini cardellini@ing.uniroma2.it, Corrado Cerruti corrado.cerruti@uniroma2.it NOVA: Mauro Castelli mcastelli@novaims.unl.pt SDU: Mikkel Baun Kjærgaard mbkj@mmmi.sdu.dk UC3M: Francisco Javier franciscojavier.prieto@uc3m.es Brunel: Ashley Braganza, ashley.braganza@brunel.ac.uk</p>
<p>ICT-15-2019 Cloud Computing (RIA)</p> <p>Deadline: 28 March 2019</p>	<p>UC3M (Coordinator): Jesús Carretero, jesus.carretero@uc3m.es, Francisco Javier franciscojavier.prieto@uc3m.es NOVA: Mauro Castelli mcastelli@novaims.unl.pt SDU: Mikkel Baun Kjærgaard mbkj@mmmi.sdu.dk UTV: Valeria Cardellini cardellini@ing.uniroma2.it, Corrado Cerruti corrado.cerruti@uniroma2.it</p>

Other activities

Topic	Overview	People involved
Privacy	<p>Plans for a YERUN workshop on privacy, data protection and ethics. Particular interest in:</p> <ul style="list-style-type: none"> • the impact of the GDPR: DCU are working on a data flow template which would map which data sets can be combined within the rules of the GDPR • Interpretability and accountability of machine learning in relation to decision making • The role of informed consent, especially around IoT and face recognition technology 	Essex SDU DCU
Education	<p>Analysing data from MOOCs to predict behaviour, results and drop-out rates in order to intervene to prevent this. Agreement to find funding for further work to pursue this, for example Erasmus+ or national funding agencies.</p>	UAM UAB UPF NOVA Maastricht
Data science methodology	<p>Planning for three more workshops in this area in:</p> <ul style="list-style-type: none"> • Madrid • Colchester • Ulm <p>A cross-sectional research theme, with particular interest in AI interpretability, mining/cherry picking of data and uncertainty.</p>	UAM - Essex – Berthold Lausen Ulm – Jan Beyersmann, Gunter Loeffler CarlosIII – Daniel Pena

PARTICIPANT BIOGRAPHIES

Prof Pedro Barahona, NOVA Lisboa

Email: pb@fct.unl.pt



Pedro Barahona is Professor of Computer Science in the Department of Computer Science (Departamento de Informática) of the Faculty of Science and Technology (Faculdade de Ciências e Tecnologia) of the New University of Lisbon (Universidade NOVA de Lisboa), where he was the Head of Department (in 1999-2006) and Director of its Centre for Artificial Intelligence - CENTRIA (in 2008-2012). His main research area is Constraint Programming, within the wider area of Artificial Intelligence, and in their applications, namely in Bioinformatics, Medicine, and Logistics, presenting regularly the result of his research in various scientific conferences and publications in these areas. To support this research, he has been responsible or participated in a number of projects, both national and international, often with a leading role. He has also organised and/or collaborated in the organisation of several scientific events, and took responsibility in the management of national and international scientific associations. In his teaching activity, he has been responsible for various courses, at graduate level as well as post-graduate, namely in Computer Science and Applied Artificial Intelligence, and supervised several M.Sc. and Ph.D. students. He is currently the Coordinator of the Master in Analysis and Engineering of Big Data and the Local Coordinator of the European Master in Computational Logic.

Prof Jan Beyersmann, Institute of Statistics, Ulm University

Email: jan.beyersmann@uni-ulm.de



Jan is Professor of Biostatistics at Ulm University and his research interests include Survival and Event History Analysis and Statistical Methodology for Clinical and Epidemiological Studies. He is Associate Editor Statistics in Medicine, Statistical Board Intensive Care Medicine, Advisory Board of the German Region of the International Biometric Society (2013-2017). He is currently working on three research projects: Combatting Bacterial Resistance in Europe (EU-funded projects COMBACTE-MAGNET and COMBACTE-CARE) - Planning and nonparametric inference for multistate time-to-event data such as disease occurrences and disease durations (German Research Foundation grant BE 4500/1-1); Drug induced adverse pregnancy outcomes: innovative event history analysis for non-continuously exposed pregnancies in the national German, and; Embryotox patient database (German Research Foundation grant BE 4500/3-1).

Dr Simone Borra, University of Rome Tor Vergata

Email: borra@economia.uniroma2.it

I am associate professor in Statistics in the Economics Faculty of University of Rome "Tor Vergata". My Postgraduate courses are: Business Statistics; Business Intelligence & Data Mining; Statistical methods for management. I am coordinator of a one year post graduate Master in Customer Experience & Social Media Analytics with the partnership of SAS Institute and Accenture. My main research interests are in: Statistical methods for Data Mining; Nonparametric methods to measure the prediction error.



Prof Ashley Braganza, Brunel University London

Email: Ashley.Braganza@brunel.ac.uk



Ashley is Deputy Dean of the College of Business, Arts and Social Sciences and Professor of Organizational Transformation at Brunel Business School. His research and consultancy expertise covers the development and implementation of large scale IT-enabled business change. His research incorporates perspectives of big data, strategy, structure, culture and behaviour, information systems, knowledge management and process reengineering. He has led several major industry-funded research projects in the area of change and process management. He completed a study of CEOs and CIOs to gain fresh insights into the role of CEOs in large scale transformation programmes. He has published three books and over a hundred research articles, conference papers and working papers covering a range of topics on big data, business processes, change management, process orientation, knowledge management, governance, and organisation structure. Prior to being appointed Deputy Dean, Ashley was Head of Economics and Finance and, as a member of the Business School's Executive Board and Advisory Board, was responsible for three major portfolios: Executive Development, Student Employability and Alumni. In addition to these roles, Ashley ran nexus – The Knowledge Exchange. nexus is a research club that examines the interplay between Change Management, Business Process, Project and Knowledge Management. nexus is a collaboration between practitioners and academics to co-create knowledge and practical ways of successfully implementing complex change programmes. Past members include BT, Network Rail, Microsoft, Astra Zeneca, Friends Provident, Volkswagen Financial Services and the Department for Children, Schools and Families. Prior to joining Brunel Business School, Ashley was with the Cranfield School of Management for seventeen years in a variety of senior academic roles. Ashley has worked with a large number of organisations advising them on a variety of business process, transformation, strategic, knowledge management and information systems issues. These projects involved multiple stakeholders with conflicting priorities and competing agendas. Ashley attained his doctorate, which focused upon the implementation of radical process oriented change, from Cranfield University. Ashley completed a MBA at the Strathclyde Business School. He elected to specialise in corporate strategy and finance. His MBA thesis was an examination of Mergers and Acquisitions in the Single European Market. Prior to his MBA, he completed the Chartered Institute of Bankers examinations. Before joining academia, Ashley spent ten years in the banking industry.

Dr Valeria Cardellini, University of Rome Tor Vergata

Email: cardellini@ing.uniroma2.it



Valeria Cardellini is associate professor in Computer Science at the University of Rome Tor Vergata, Italy. Her main research interests focus on the design, engineering and performance evaluation of distributed systems and applications. In the last years she has been working on QoS-driven adaptation of service oriented applications, resource provisioning and pricing in Cloud systems, computation offloading in mobile Cloud computing. In the Big Data realm she focuses on the second V (Velocity) of the triad, investigating how to

autonomously manage and adapt at runtime Data Stream Processing applications in a geo-distributed environment as envisioned by the Fog computing paradigm. She has published more than 90 referred papers in international journals, book chapters, and conference proceedings and edited two books. Three of these publications have received paper awards. She contributes and contributed actively in several international and national research projects, including the EU Cost Action ACROSS and EoCoE center of excellence. She holds a PhD in Computer Science from the University of Rome Tor Vergata for her work on scalable Web-server systems.

Prof Mauro Castelli, NOVA Lisboa

Email: mcastelli@novaims.unl.pt

Mauro Castelli received the master's (summa cum laude) degree in computer science and the Ph.D. degree from the University of Milano Bicocca, Milan, Italy, in 2008 and 2012, respectively. He is an Assistant Professor with the NOVA IMS, Universidade Nova de Lisboa, Lisbon, Portugal. His current research interests include artificial intelligence, in particular evolutionary computation and genetic programming, and in the application of machine learning techniques to solve complex real-life problems, especially in the field of biology and medicine.



Dr Corrado Cerruti, University of Rome Tor Vergata

Email: corrado.cerruti@uniroma2.it



Corrado is full professor in innovation management and his main research interests focus on the impact of digital transformation and business processes and business models with a specific reference to the area of supply chain management and the industries of Management Consulting and Utilities. More specifically he is presently coordinating a research project funded by FEACO (Federation of European Association of Management Consulting) examining how management consulting practices are changing as a result of digital transformation and whether the current value proposition and business model of mainstream consulting companies are viable in the digital economy. He is director of the MSc in Business Administration that includes a specialization in Management Consulting (mainly oriented towards digital transformation) and he is the coordinator of the summer school on "Management Consulting in the Digital Age" (hosting more than 20 students from 8 YERUN universities). The research and teaching activities are carried out in cooperation with leading IT vendors and consulting companies such as SAS, SAP, Accenture, Cap Gemini and PwC.

Dr Ignacio Criado Grande, Universidad Autonoma de Madrid

Email: ignacio.criado@uam.es



Associate professor at Department of Political Science and International Relations, Universidad Autónoma de Madrid, Spain and research fellow at Center for Technology in Government, State University of New York (SUNY at Albany). I am an active researcher in the field of social media technologies in public policy and administration, open government and open data policies. During the last years, I have delivered research on social media data analysis, with a special focus on local governments, public

services and public policy. Besides, I am very interested in the policy process and governance of public open data, big data analytics, and social data science. I work using emergent methods, including Social Network Analysis (SNA) and Sentiment Analysis. Currently, I am principal investigator (PI) of the BBVA Foundation project: Smart Governance of Digital Social Media in Local Governments within the European Union. Implementation Strategies, Success Factors and Public Service Models in Spain, the Netherlands, the United Kingdom and Sweden.

Dr Paola Di Giuseppantonio Di Franco, University of Essex

Email: pd17425@essex.ac.uk

Paola is Eastern ARC Fellow for Digital Humanities at the University of Essex. My research combines material culture, heritage, and cognitive science to explore how new technologies impact heritage making processes and the interpretation processes of the past. My recent Marie Skłodowska Curie project, titled DIGIFACT: Digital artefacts: How People Perceive Tangible Cultural Heritage through Different Media has clarified the role of new technologies in the perception and understanding of cultural heritage; specifically how 3D virtual and material replicas can re-define museum practices. This project was in collaboration with the Museum of Anthropology and Archaeology in Cambridge, developing a research programme to feed into the redevelopment of the World Archaeology Gallery over the next years. I am now developing two new projects that will further her line of research aimed at analysing how new technologies impact heritage making and interpretation processes of the past: one project is investigating the role new technologies might play in the rebuilding process that societies face after suffering natural disasters or environmental trauma. The second project combines education, history and heritage for the digital preservation of tangible and intangible aspects of heritage associated with the 19th cent. Palace of Said in Tunis. The project is a collaboration with the faculty of Education, University of Cambridge, Virtual Experience company, and Rambourg Foundation in Tunis and it is supported by the British Council's Cultural Protection Fund, which aims to preserve cultural artefacts in conflict-affected countries. During the project the team has produced an interactive 3D model of the Palace of Said and developed learning resources to build Tunisian students' critical engagement with their heritage.



Prof Maria Fasli, University of Essex

Email: mfasli@essex.ac.uk



Maria is Director the Institute for Analytics and Data Science at the University of Essex and the first UNESCO Chair in Analytics and Data Science. She also leads the Methodological Stream for Big Data Analytics in the ESRC BLG Data Research Centre. Her area of expertise is in artificial intelligence and smart analytics. She served as Head of School of Computer Science and Electronic Engineering between 8/2009-12/2014 and has demonstrated successful leadership evidenced by discernible and substantial outcomes in support of the University's strategic priorities. She has been instrumental in building strategic partnerships internationally and developing joint programmes with partners (China), and forging links with businesses and other organisations. She has been developing and delivering innovative programmes and supervising students (9 PhDs completed; currently supervising 7). In 2005, she was

awarded a National Teaching Fellowship by the Higher Education Academy (UK), for her innovative approaches to education. She has organised international workshops and events in both research and education.

Prof Anibal R Figueiras-Vidal, Universidad Carlos III de Madrid

Email: arfv@tsc.uc3m.es

Full Professor Universidad Carlos III de Madrid, with research interests in: Machine Learning, mainly - deep learning- imbalanced classification and applications, mainly- imbalanced problems (fraud, failures, security & safety, diagnostics.), smart cities & environments-health- business.



Prof Daniel Franco Puntos, University Autonoma of Barcelona

Email: daniel.franco@uab.cat / dr.escola.enginyeria@uab.cat



Dr. Daniel Franco is currently Dean of the Engineering School. Formerly, he has been programme coordinator and vice-dean. He has received a Bs. in Computer Science in 1991, Ms. in Science in Computer Science in 1993 and a PhD in Computer Science by the UAB in 2000. From then, he has been an associate professor at Department of Computer Architecture and Operating Systems at the Engineering School of the University Autonoma of Barcelona. He has been teaching at Bachelor and Master levels in Computer Science, Telecommunications, and Electronic Engineering programs, among others. He has implemented several teaching innovations and has some publications about this topic, as well as about quality management in education. His research activities have been focused in the fields of Parallel Computing, Computer Architecture, Computer Simulation, etc. being always involved in competitive research projects (national and international), and he has been advisor of 3 doctoral theses and several master theses. He has more than 20 publications both in top conferences and journals of his research area; he has collaborated with the industry and holds a registered patent. He has been awarded with two 6-year national research evaluation periods.

His research interests are focused on: autonomous and connected vehicles; computer vision; sensing; opportunistic communications; GPUs; embedded /FPGA /Ultra low power; AI /deep learning; cloud systems /cloud and distributed systems management, and; cybersecurity /blockchain.

Prof Bart Goethals, University of Antwerp

Email: bart.goethals@uantwerpen.be



Bart Goethals is professor at the Department of Mathematics and Computer Science of the University of Antwerp in Belgium where he leads the Adrem Data Lab, which performs fundamental research in Data Science. His primary research interests are the study of data mining techniques to efficiently find interesting patterns and properties in large databases. He received several awards, chaired some of the most influential academic Data Science conferences, and serves on the editorial board of three highly ranked scientific journals. In 2016 he co-founded the university spin-off company Froomle, whose goal is to boost its customer's sales and conversion by making their consumer experiences more relevant and personalised using artificial intelligence.

Mr Sebastian Haunss, University of Bremen

Email: sebastian.haunss@uni-bremen.de

Sebastian Haunss is a senior researcher in political science and leads a research group on social conflicts at the Research Center on Inequality and Social Policy at the University of Bremen. His research interests are global social policy, social conflicts and political mobilizations in the knowledge society, social networks and social movements. Sebastian Haunss is currently leading two research projects. In the project MARDY (Modeling ARGumentation DYNamics in Political Discourse) an interdisciplinary team from Political Science, Computational Linguistics and Machine Learning aims to develop a framework for data-driven modeling of key aspects of argumentation dynamics in policy debates as they unfold over a period of days or weeks. The project "Global developments of health systems and long-term care as a new social risk" a research team from political science and health economics will analyze the development of health systems and long-term care in a global perspective. The aim is to create a limited typology and to explain to what extent similarities and differences between national systems with horizontal social, economic and cultural exchange relations between countries as well as through vertical exchange relations between countries and international organizations.



Dr Markus Helfert, Dublin City University

Email: markus.helfert@dcu.ie

Dr. Markus Helfert is Director of the Business Informatics Group at Dublin City University, and an Associate Professor in Information Systems at the School of Computing, Dublin City University (Ireland). He is a Science Foundation Ireland funded Investigator at Insight – the Centre for Data Analytics. His research interests include Data Management, Data Governance, Big Data, Cloud Computing, Service Science, Enterprise Architecture and Open Data and Smart Cities. Dr. Helfert has received national and international grants from agencies such as European Union (FP7; H2020), Science Foundation Ireland and Enterprise Ireland. Markus Helfert is co-ordinator of PERFORM – A Marie Skłodowska-Curie training network for young researchers in



the field of Digital Retail. He has authored more than 200+ academic articles, journal and book contributions and has presented his work at international conferences.

Prof Roberto Henriques, NOVA Lisboa

Email: roberto@novaims.unl.pt



Roberto Henriques is currently Assistant Professor at NOVA Information Management School (NOVA IMS) where is the coordinator of the Masters in Information Management and Statistics, in Information Management and the European Master of Science in Information Systems Management. He holds a Ph.D. in Information Management (NOVA), a Masters in Geographic Information Systems and Science and a degree in Biophysics Engineering. His research interest includes geospatial data analysis and decision support systems using artificial intelligence and data mining methods. His work has been published in several international high-quality journals and conferences. He has also been involved in several projects including marketing analytics and CRM, client management and data mining.

Mr Mikkel Kjærgaard, University of Southern Denmark

Email: mbkj@mmmi.sdu.dk



Mikkel Baun Kjærgaard conducts research within the area of mobile and pervasive sensing systems with a specific focus on energy informatics, occupancy behavior, crowd behavior, positioning and machine learning. He conducts research with an experimental foundation at the intersection of ubiquitous computing, machine learning and systems research. His work on positioning covers contributions advancing the method of Location Fingerprinting for indoor positioning, and middleware for positioning and sensor fusion using inertial sensors. He has contributed with methods to address power consumption issues for mobile sensing, methods for collective sensing and visualisation focusing on the detection of crowd behaviors, and sensor-based understanding of occupancy behavior and electricity usage.

Prof Mikko Laitinen, University of Eastern Finland

Email: mikko.laitinen@uef.fi



Mikko Laitinen is Professor of English Language at the University of Eastern Finland. He was previously Professor of English Linguistics at Linnaeus University, where he is one of the two founding members of Data-Intensive Sciences and Applications (DISA), a multidisciplinary research consortium involving computer scientists, mathematicians, and social scientists. His research interests include corpus linguistics and digital tools in linguistics and variationist approaches to English as a lingua franca (ELF). Laitinen's research focuses on understanding language contact and emerging variability in the digital age. It seeks to combine traditional philological data with real-time data and to develop and apply new methods to enrich and visually analyze natural language data streams that are large in scope and rich in metadata to gain novel insights on language variability in social context. He has published nearly 50 peer-reviewed

articles including contributions in Language Variation and Change, International Journal of Corpus Linguistics, International Journal of the Sociology of Language, and Journal of Universal Computing (forthcoming). Laitinen has worked in several language corpora projects. He is currently working with building the first second-generation ELF corpora, the Corpus of Written English Texts in Sweden/Finland (SWE-CE and FIN-CE), and is developing tools for new third-generation ELF corpora that consist of real-time data streams of text and metadata from social media platforms, such as Twitter. His recent publication in this area is a study that focuses on the weak social tie model of linguistic change. This model has been tested on very small ethnographic networks, and this study operationalizes weak ties using circa 200,000 accounts and 10 million individual tweets and their automatically-generated metadata parameters to test whether the weak tie model also holds in big data. It serves to illustrate what type of new answers to traditional questions could be gained by using new types of 'big data' evidence in humanities.

Prof Berthold Lausen, University of Essex

Email: blausen@essex.ac.uk



President (2018-19) of the International Federation of Classification Societies (IFCS), president of the Gesellschaft fuer Klassifikation (GfKI) Data Science Society and vice president of the European Association for Data Science (EuADS). As head of the department (2016-) and Professor of Data Science he is affiliated to the Department of Mathematical Sciences, University of Essex, Colchester, UK. At

Essex he lead the introduction of a new Master of Data Science in 2014 which is delivered by the Department of Mathematical Sciences (DMS) and the School of Computer Science and Electronic Engineering (SCSEE). He is principal investigator of a knowledge transfer partnership funded by Innovate UK between the University of Essex and the data science company Profusion in London. He contributes to the ESRC Business and Local Government Data Research Centre at Essex and supports as medical statistician the NIHR Research Design Service for the East of England. He is a member of the Institute for Analytics and Data Science (IADS) and the Essex Biomedical Sciences Institute (EBSI) and an external member of the Medical School, University of Erlangen-Nuremberg. He studied statistics at the Faculty of Statistics, Technical University of Dortmund (Diplom 1987; Dr. rer. nat. 1990).

Prof Gunter Löffler, Ulm University

Email: gunter.loeffler@uni-ulm.de



Gunter Löffler is professor of finance at Ulm University. His current research interests are on risk management and empirical finance, with a focus on predictive models. Gunter is co-author of the book "Credit Risk Modeling using Excel and VBA". He was assistant professor at Goethe University Frankfurt and served as an internal consultant in the asset management division of Commerzbank. His PhD in finance is from the University of Mannheim. Gunter has studied at Heidelberg and Cambridge Universities.

Dr Miguel A. Mayer, Universitat Pompeu Fabra

Email: miquelangel.mayer@upf.edu



Miguel A. Mayer, MD, PhD, MPH (IMIM-UPF, Barcelona, Spain) is a Specialist in Family and Community Medicine. For many years he worked as a clinician for the Spanish National Health Service in an Internal Medicine Department and in Primary Care settings. He was the director of the Department of Health Web Site Accreditation Program of the Medical Association of Barcelona, being in charge of a leading international quality program of health-related content websites. He is a Senior Data Scientist at the Research Programme on Biomedical Informatics (GRIB) of

the Hospital del Mar Medical Research Institute (IMIM) and Associate Professor and academic coordinator of the Biomedical Informatics subject at the Universitat Pompeu Fabra (UPF) Faculty of Health and Life Sciences in Barcelona, Spain. He is and has been involved in several funded European projects since 2002, within programs such as the Safer Internet Action Plan, DG SANCO, Horizon2020 and the Innovative Medicines Initiative (IMI), working on several research topics such as new health ICT strategies, quality assessment of health information on the Internet, Ethics and public health, standards in the medical domain, the reuse of EHR for research, as well as the analysis of Web 2.0 and mobile Apps for medical and public health applications, for better access and management of health big data information. He is a member of several scientific working groups such as the Big Data Value Association (BDVA) of the EU Commission and the Participatory Health & Social Media WG of the International Medical Informatics Association (IMIA).

Dr Dimitri Ognibene, University of Essex

Email: dimitri.ognibene@essex.ac.uk



Dimitri Ognibene obtained his PhD in Robotics from the University of Genoa in 2009. Before joining the University of Essex, he has been performing experimental studies and developing formal methods for active social perception at UPF, Barcelona, as a Marie Skłodowska-Curie COFUND Fellow; developing algorithms for active vision in industrial robotic tasks as a Research Associate (RA) at Centre for Robotics Research, Kings College London; devising Bayesian methods and robotic models for

attention in social and dynamic environments as a RA at the Personal Robotics Laboratory in Imperial College London; studying interaction between active vision and autonomous learning in neuro-robotic models as a RA at Institute of Cognitive Science and Technologies of the Italian Research Council (ISTC CNR). He also collaborated with Wellcome Trust Centre for Neuroimaging (UCL) to address the exploration issue in the currently dominant neurocomputational modelling paradigm. Dr Ognibene has also been Visiting Researcher at Bounded Resource Reasoning Laboratory in UMass and at University of Reykjavik (Iceland) exploring the symmetries between active sensor control and active computation or metareasoning. Dr Ognibene presented his work in several international conferences on artificial intelligence, adaptation, and development and published on international peer-reviewed journals. Dr Ognibene was invited to speak at the International Symposium for Attention in Cognitive Systems (2013 and 2014) as well as in other various neuroscience, robotics and machine-learning international venues. Dr Ognibene is Associate Editor of *Paladyn, Journal of Behavioral Robotics*, and has been part of

the Program Committee of several conferences and symposiums. His research interests entail the principles underlying adaptive social behaviour in complex conditions with particular attention to anticipative skills and exploration in unknown environments. His research covers both biological and artificial agents as well as physical (visual and multimodal) and virtual environments as social media and virtual reality. He uses machine learning methods, both (deep) neural and probabilistic models, to create agents as well as to study and predict their behaviours.

Prof Daniel Peña, Universidad Carlos III de Madrid

Email: daniel.pena@uc3m.es

Professor of Statistics and Director of the Financial Big Data Institute at Universidad Carlos III de Madrid. His research interests include Statistics and Machine learning for Big Data. Time series, dimension reduction, clustering.



Dr Estrella Pulido, Universidad Autonoma de Madrid

Email: estrella.pulido@uam.es



Dr. Estrella Pulido obtained a degree in Computer Science in 1989 from the Computer Science Faculty of the Universidad Politécnica de Madrid. I was working in industry from 1987 until 1991 in R & D departments. I obtained a MSc in Artificial Intelligence with honours from the University of Bristol in 1992 and was granted a Faculty fellowship from the same University to carry out a PhD which I finalised in 1996. Currently I work at the Escuela Politécnica Superior of the Universidad Autónoma de Madrid (UAM) since October 1996 where I hold an associate professor position since July 2000. My research focuses on two main fields: the use of ICT in Education and big data and data analysis. In both fields I have participated in different projects financed by public calls both nationally and internationally and I am the author of more than fifty scientific-technical publications. Among these projects, I would like to mention the project "Frontiers in Automatic Learning and Multidisciplinary Applications" in which the two mentioned fields converge. My responsibility in this project is the application of analysis techniques to the data collected in the learning process that allow understanding and predicting the behaviour of students in online teaching environments to optimize their learning process. I am co-director of the UAM / IBM Chair in Mainframes, Cognitive Solutions and Big Data Analytics, signed in 2013 whose main objective is the promotion of teaching, research and dissemination of knowledge in mainframes, cognitive computing and big data. As part of the activities of the Chair, I have directed and coordinated several training courses on Big Data and Hadoop technology as well as on the use of Mainframes. In this line of training, I am director of the Master in Big Data and Data Science of the UAM whose main goal is to train students in the most relevant topics on Big Data that have to do with the collection, storage and processing of data (Hadoop, Spark, etc.), the infrastructure for Big Data (multicore, security), data analysis (machine learning) and visualization. I have coordinated and participated in different teaching innovation projects through calls funded by the UAM

and I am currently part of a team of teachers that coordinate a MOOC on Android Application Programming offered through the edX online teaching platform. In terms of management tasks, I held the positions of Rector's Delegate for Network Information from 1999 to 2002, one of my main functions being technical advice for the design, implementation and maintenance of the UAM website, Director's Delegate for the Website and Institutional Image of the Higher Polytechnics School from 2005 to 2008, responsible for the design, management and maintenance of the web portal of the mentioned School's website. I was the Director of the Higher Polytechnic School from 2008 to 2012.

Mr Juan Romo Urroz, Universidad Carlos III de Madrid

Email: rector@uc3m.es

Rector of Universidad Carlos III de Madrid and President of YERUN. Research interests are complex and high-dimensional data analysis (including functional data analysis, big data, resampling, time series, extremes theory and outliers) and applications (including microarrays in genetics, geophysics, financial data and image analysis).



Dr Victor Rubio, Associate Professor, Universidad Autonoma de Madrid

Email: victor.rubio@uam.es



My background is in Health Psychology and I have always been keen on how to assess the different dimensions relevant to the phenomena we study or counsel. I'm currently interested in the field of sport and exercise, particularly in preventing sport injuries. I have recently launch a project aimed at testing whether big data analysis based on body motion and facial expression might accurately produce individuals' personality and other psychological variables judgments.

Dr Spyros Samothrakis, University of Essex

Email: ssamot@essex.ac.uk



Spyros is Research Fellow and Assistant Director of the Institute for Data Analytics at the University of Essex. My research interests include Reinforcement Learning (with a current emphasis in incremental learning) for conversational and game playing agents, as well the causal analysis of observational data.

Prof Peter Schneider-Kamp, Professor, University of Southern Denmark

Email: petersk@imada.sdu.dk



Today we live in a world where software systems and data are ubiquitous - from the super computers predicting our climate to the very mobile phones in our pockets. The complexity of these ICT systems is continuously growing as even mobile phones approach the power and versatility of standard office computers from 5 years ago. Consequently, building functional, reliable, and efficient ICT systems has become one of the main challenges in the development of new products and services. I strongly believe that the growing complexity of ICT systems development can only be handled by supporting the development process with more powerful software tools, where possible based on strong mathematical principles. My research background and my core research interests are in automated reasoning, addressing foundational issues in computer science and mathematics as well as a wide range of applications. Most of my results so far have been in automated software verification, automated synthesis, and automated distributed programming. Since 2016 I have been heading the strategic research focus area “Open Data Exploration” (ODEx™), which currently consists of 10 interdisciplinary projects ranging from smart traffic systems to data-driven life cycle economics. Currently I am starting up newly granted research projects on automated anonymization of sensitive data in the wake of the soon arriving EU GDPR and on automated drone inspection of power lines.

Dr Evgueni Smirnov, Maastricht University

Email: smirnov@maastrichtuniversity.nl



Evgueni Smirnov is an assistant professor of Artificial Intelligence at the Department of Data Science and Knowledge Engineering, Maastricht University. His research interests include: Data mining (reliable prediction, association-rule mining); Machine learning (ensemble learning, online learning, kernel methods, transfer learning, version spaces); Recommender systems; Applications of machine learning in medicine, education, and transportation. Evgueni Smirnov gives machine-learning related courses and computer-science related courses for the Bachelor program in Data Science and Knowledge Engineering and the Master program in Data Science of Maastricht University. Evgueni Smirnov supervised/executed seven commercial data-mining projects. His team of PhD students does research on ensemble learning, transfer learning, and medical data mining.

Prof Doroteo Torre Toledano, Universidad Autonoma de Madrid

Email: doroteo.torre@uam.es



Prof. Doroteo Torre Toledano, received the M.S. degree in 1997, and the Ph.D. degree Electrical and Electronic Engineering in 2001, both from Universidad Politecnica de Madrid, Spain. He has been recipient of several academic awards, such as the First National Bachelor Award of Spain, the best academic record in Electrical and Electronic Engineering (of 448 students) and a Ph.D. Dissertation Award from the Spanish Association of Telecommunication Engineers. After his Ph.D., he joined M.I.T. as Postdoctoral Research Associate in the Spoken Language Systems Group (2001-2002), under the supervision of Profs. Victor Zue and Jim Glass. He has also experience working in the industry, in particular in the Speech Technology Division of Telefonica R&D, where he worked from 1994 to 2001 and also in 2003. His trajectory as professor in signal processing starts in 2004, when he joined Universidad Autonoma de Madrid, where he is currently Associate Professor and has been recently accredited as Full Professor. Prof. Toledano has over 20 years of experience in speech processing, over 100 scientific publications that have received 1400 cites according to Google Scholar and an h-index of 18. He has participated in 6 EU research projects and in over 40 national projects (in 10 of them as principal investigator). He has participated in over 15 technological competitive evaluations (mainly NIST evaluations) and has organized three. He was General Co-Chair and main organizer of IberSPEECH 2012, and organizer and session chair of several other conferences. Prof. Toledano current research is focused on speech, speaker, language and pathology recognition, particularly based on deep learning approaches.

Workshop Facilitator

Carol Wilson, Co-Founder, Wilson Sherriff



Carol Sherriff is a Certified Professional Facilitator™, coach and consultant with more than 15 years' experience working in Europe, North America and Asia. She has worked extensively with higher education institutions, research councils and UN bodies facilitating collaborative approaches to research and innovation. She also holds academic roles at the Open University Business School and University of Hertfordshire and is a Fellow of the Higher Education Academy.