

Chalmers Initiative Seminar on Big Data

25th – 26th March 2014



Welcome!



David Sands

**Director of
Chalmers ICT
Area of Advance**

We are surrounded by data. Every day, sensors around us and in our digital lives generate terabytes of data based on our activities. Our bodies are a data source too, with billions of data sequences found in our genetic codes. Our business activities are permeated with data collection. How do we deal with such a large influx of data? How do we benefit from the new opportunities created by large-scale data sets? How do we find the human face of big data?

Initiative Seminars at Chalmers represents crucial meeting places for people representing research, innovation and society. The seminars address challenging questions for humanity, important for a sustainable future. I hope that this Initiative Seminar on Big Data, hosted together with the University of Gothenburg, will be only the beginning of an in-depth discussion on our data-oriented future. Our aim is to inspire discussions and research initiatives that will enable data-driven innovation in Sweden and beyond. I wish you a pleasant experience, immersing discussions, and a lot to think about.

**Welcome to Chalmers
Initiative Seminar on Big Data!**

Day 1 – Before lunch Tuesday 25th March 2014

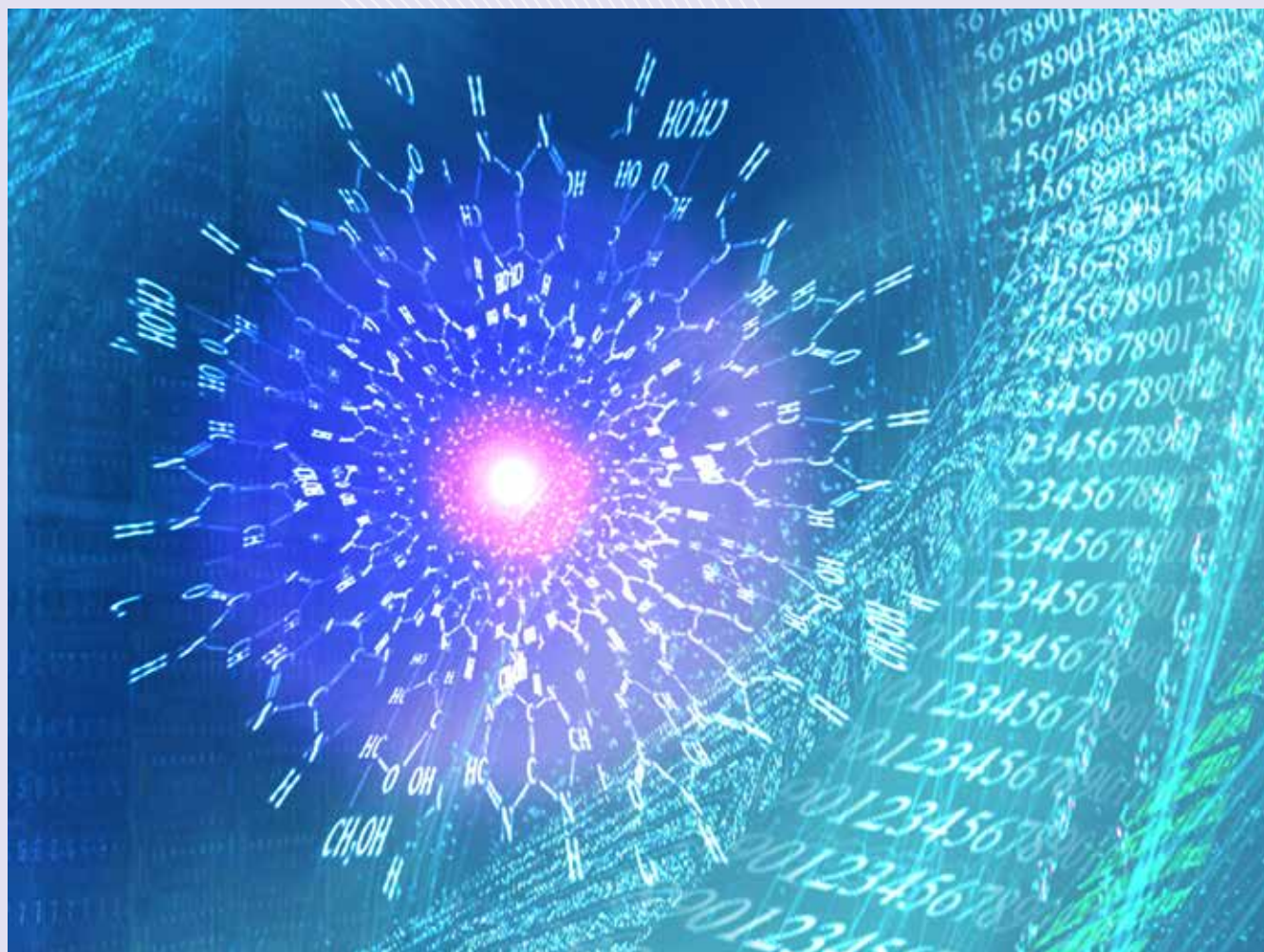


Illustration: Simon Fletcher

Programme

08:30 – Registration and coffee

09:00 – Welcome and overview of the first day

Big Data in the Life Sciences and Medicine. Session chair: Devdatt Dubhashi

09:10 – Terry Speed, Walter & Eliza Hall Institute of Medical Research, Australia. UC Berkeley, USA. “Data Science, Big Data and Statistics: can we all live together?”

10:00 – Janusz Dutkowski, UC San Diego, USA. “From Networks to Ontologies of Gene Function.”

10:30 – Coffee, sandwich, exhibition and posters

11:00 – Helen Seeman-Lodding, eHealth department, HSA, Region Västra Götaland, Sweden. “Will we dare to release Healthcare data?”

11:10 – Mikael Wintell, Lars Lindsköld, eHealth department, HSA, Region Västra Götaland, Sweden. “The VGR way – to handle Big that will get Bigger? Experience after 8 year and many Terabytes”

11:35 – Göran Bergström, SCAPIS, Sahlgrenska Universitetssjukhuset, Region Västra Götaland, Sweden. “It’s all about the heart”

12:00 – Hans Enocson, GE Healthcare/Global Operations, Sweden. “How Big Data creates opportunities for productivity improvements”

12:30 – Lunch, exhibition and posters



Data Science, Big Data and Statistics: can we all live together?
Terry Speed

Terry Speed will report some reflections on these issues, offer some suggestions for statisticians, and summarise some theory which, in his opinion, has relevance to the analysis of data, whoever does it.

Terry Speed's research interests lie in the application of statistics to genetics and genomics, and to related fields such as proteomics, metabolomics and epigenomics. He works at the Walter & Eliza Hall Institute of Medical Research in Melbourne, and is an active emeritus professor in Statistics at UC Berkeley.



Will we dare to release Healthcare data?
Helen Seeman-Lodding

Will we dare to release the power of Healthcare Data, were the information can be weightlessly transferred around the globe and by it shelf try to find cure for its sake.

Helen Seeman-Lodding MD., Ph.D., Senior Consultant Anaesthesiology and Intensive Care Medicine, eHealth department, HSA, Region Västra Götaland Sweden. Apart from active clinical work and research tutoring, Helen is also responsible for a development program regarding future digital systems of clinical information management in Region Västra Götaland Sweden.



How Big Data creates opportunities for productivity improvements
Hans Enocson

An overview of how Big Data creates opportunities for productivity improvements, with examples from the business and health care sectors. GE has worked with Industrial Internet during several years, and sees this as a very important component in developing new service concepts and solutions for their customers.

Hans Enocson is President & CEO of GE for the Nordic and Baltic Region (Estonia and Latvia) since 2007. He is primarily responsible for GE's overall growth in the region and acts as the key interface to GE for governmental and corporate relations. Hans has more than 30 years of industrial experience and has served internationally in global responsibilities, working abroad in France and USA. He serves on many Nordic GE boards, AmCham Sweden, SAAC NY and Uppsala Innovation Center. From academic perspective he holds various advisory roles to the University of Uppsala and Stockholm Resilience Center. He is a frequent speaker on topics such as Sustainability, Corporate Governance and Innovation. Hans holds a BA from the University of Stockholm and Program for Executive Development (PED) from IMD Lausanne, Switzerland.



From Networks to Ontologies of Gene Function
Janusz Dutkowski

Dr. Dutkowski will talk about our efforts to turn complex networks of gene and protein interactions into an ontology of biological processes and functions. He will compare our data-driven ontology to an ontology constructed manually through literature curation.

Dr. Janusz Dutkowski is a project scientist at the UCSD School of Medicine. His research interests fall broadly within the field of computational biology with emphasis on new approaches for extracting knowledge from biological networks and large cancer genomics datasets.



It's all about the Heart
Göran Bergström

Presentation of the Swedish CARDioPulmonary biolmage Study (SCAPIS) which aim to build a nationwide, open-access, population-based cohort for the study of cardiovascular disease (CVD) and chronic obstructive pulmonary disease (COPD). SCAPIS will recruit and investigate 30,000 subjects with detailed imaging and analyses of the cardiovascular and pulmonary systems.

Göran Bergström is head of the Physiology Group at Wallenberg Laboratory and senior consultant in clinical physiology at the Vascular Diagnostic Unit, Sahlgrenska University Hospital. He is currently leading the pilot trial of SCAPIS (Swedish CARDio-Pulmonary biolmage Study), funded by the Swedish Heart-Lung foundation. Professor Bergström's research focus is on the vulnerable atherosclerotic plaque. His group uses the atherosclerotic carotid artery as an experimental model to study the pathophysiology, diagnostics and epidemiology of atherosclerosis.



The VGR way – to handle Big that will get Bigger?
Mikael Wintell



Lars Lindsköld

Will you continue to do as you have always done, or should you do something you have never done – in order to achieve greater Value? That is the "BIG" question.

Mikael Wintell, Chairman DICOM WG 26, Member IHE, HL7, ISO TC215, eHealth department, HSA, Region Västra Götaland Sweden. Long experience in medical imaging, started to work with radiology 1989 in the Goteborg area. Responsible for the development of CT information exchange to gain greater knowledge cooperation. CMIO of the largest repository of diagnostic medical data in region Västra Götaland, Sweden.

Lars Lindsköld Ph.D., eHealth department, HSA, Region Västra Götaland Sweden. CEO - of the largest repository of radiology information (images and reports combined) in Region Västra Götaland, Sweden. Experience in Designing and Using an Information Infrastructure in Radiology. My focus is to create value by understanding and implement methods for extracting information from this Mimisbrunnr of wisdom.



Illustration: Simon Fletcher

Programme

Big Data Visualization and Prediction. Session chair:
 Morten Fjeld, Chalmers

14:00– Danyel Fisher, Microsoft Research, USA.
 “Design Constraints for Visualizing Big Data”

14:30 – Anders Ynnerman, Linköping University, Sweden.
 “Visual Data Analytics – The Power 2 C”

15:00 – Staffan Truvé, Recorded Future, Sweden
 “Web Intelligence for Better Business Decisions”

15:30 – Coffee, pastries, exhibition and posters

16:00 – Professor David De Roure, Oxford e-Research Centre, United Kingdom “Big Data and Social Machines”

16:30 – Panel Discussion. “Privacy, Security, and Safety in Big Data”
 Moderator: Josh COWLS, Oxford Internet Institute, United Kingdom.
 – Klas Eckervald, Teradata, Sweden
 – Helen Seeman-Lodding, Region Västra Götaland, Sweden
 – Simone Fischer-Hübner, Dept. of Computer Science Karlstad Univ., Sweden
 – Sven Ekholm, MedTech West, Sweden
 – Robin Brouwer, Astra Zeneca, Sweden

17:15 – Closing remarks



Visual Data Analytics – The Power 2C
 Anders Ynnerman

This talk will take the power of the human visual perception system as a starting point and show how computers and humans together can solve complex problems involving large-scale data. Examples will be taken from several applications including the latest data challenges in the medical domain.

Professor Anders Ynnerman received a Ph.D. in physics from University of Gothenburg. During the early 90s he was at Oxford University, UK, and Vanderbilt University, USA. From 1997 to 2002 he directed the Swedish National Supercomputer Centre and from 2002 to 2006 he directed the Swedish National Infrastructure for Computing (SNIC). Since 1999 he has held a chair in scientific visualization and is the director of the Norrköping Visualization Center – C.



**Panel Discussion:
Privacy, Security,
and Safety in
Big Data**
Josh Cowsl

Josh Cowsl researches the uses and abuses of big data at the Oxford Internet Institute. As part of a Sloan Foundation-funded project, Josh and his colleagues have interviewed practitioners pioneering the use of big data in research, generating broader insights into the opportunities and challenges big data offers for academia, business and governance



**Panel Discussion:
Privacy, Security,
and Safety in
Big Data**
Sven Ekholm

Sven Ekholm, MD 1973, PhD 1984, is a radiologist trained at the University of Lund. He moved to USA in 1981 where he was employed as a clinician and researcher in neuroradiology until 1987 at the University of Rochester, NY (UofR). In 1987 he moved back to Sweden and joined the Sahlgrenska University Hospital (SU) 1987-2003; during 1997-2003 as Professor in Radiology. After that and until 2013 he again joined the Department of Radiology at UofR, NY, USA where he stayed until he retired last summer. Shortly after his US retirement he was employed as Professor and Senior Advisor at University of Gothenburg to support the building of an academic environment within the new image and interventional center (BoIC) under construction at SU. As such he has his location within MedTech West which is closely affiliated with this new imaging endeavor.



**Panel Discussion:
Privacy, Security,
and Safety in
Big Data**
Simone
Fischer-Hübner

Simone Fischer-Hübner has been a Full Professor in Computer Science at Karlstad University since June 2000. Her research interests include IT-security, privacy and privacy-enhancing technologies. She is currently involved in interdisciplinary privacy-related research EU FP7 projects such as A4Cloud and SmartSociety. She is also member of the advisory board the Swedish Data Protection Commissioner and MSB's Information Security Advisory Board.



**Big Data and
Social Machines**
David De Roure

Will you continue to do as you have always done, or should you do something you have never done – in order to achieve greater Value? That is the “BIG” question.

David De Roure is Professor of e-Research and Director of the interdisciplinary e-Research Centre at the University of Oxford. He works closely with multiple disciplines across the social sciences (as a strategic advisor to the UK Economic and Social Science Research Council), humanities (responsible for Digital Humanities in the Oxford Research Centre for the Humanities) and sciences (astrophysics, and previously chemistry and bioinformatics). He focuses on advancing digital scholarship and new forms of scholarly communication in the context of methodological change. His personal research is in Web Science, especially “social machines”, and in computational musicology.



**Web Intelligence for
Better Business
Decisions**
Staffan Truvé

Will we dare to release the power of Healthcare Data, were the information can be weightlessly transferred around the globe and by it shelf try to find cure for its sake.

Staffan Truvé is Chief Technology Officer and co-founder of Recorded Future. Staffan has spent the last 20 years working in the borderland between research and industry, creating new companies based on cutting-edge research. He helped launch more than 15 companies, including Spotfire, Appgate, Axiomatics, Peerialism, Make-wave, Enmesh, and Recorded Future. From 2005 to 2009, he was CEO of the Swedish Institute of Computer Science (SICS) and Interactive Institute, overseeing over 200 researchers, and is today still engaged as chairman of Interactive Institute. Staffan is a member of the advisory board to the Swedish Minister of Information Technology. Staffan holds a PhD in Computer Science from Chalmers University of Technology and an MBA from University of Gothenburg. He has been a Fulbright Visiting Scholar at MIT. He is a member of the Royal Swedish Academy of Engineering Sciences.



**Design Constraints for
Visualizing Big Data**
Danyel Fisher

While big data analytics is an increasingly-popular area, we are still learning about how to conduct exploratory data analytics on big datasets. We summarize major design challenges in visualizing Big Data with four points:

- There's too much to process;
- You don't get to see it all;
- The rules are different; and
- Completely new challenges emerge

We discuss these challenges, and offer solutions to these problems.

Danyel Fisher is a researcher in information visualization and human-computer interaction at Microsoft Research's VIBE group. His research focuses on ways to help users interact with data more easily. His recent work has looked at ways to make big data analytics faster and more interactive with incremental visualization; his papers Trust Me, I'm Partially Right and Interactions with Big Data Analytics outline the research direction. Outside Microsoft, he has helped organize the “Industry and Government” track at IEEE Info Vis Conference, bringing together practitioners with academics at the premier visualization conference. Danyel received his MS from UC Berkeley, and his PhD from UC Irvine.



**Panel Discussion:
Privacy, Security,
and Safety in
Big Data**
Robin Brouwer

As R&D Architecture Practice Leader at Astra Zeneca, Robin Brouwer is accountable for the global architecture capability development. Over the past two years the capability of Astra Zeneca has transformed towards one that is driving for business impact with a strong technology skills backbone.



Klas Eckervald

Klas Eckervald (at Teradata) helps leading European organizations getting value out of Data by leveraging the latest advances in DW research.

Day 2 – Before lunch

Wednesday 26th March 2014



Illustration: Simon Fletcher

Programme

08:30 – Registration and coffee

09:00 – Welcome and overview of the second day

09:10 – Bo Rothstein, Department of Political Science, University of Gothenburg, Sweden. “Quality of Government and Human Well Being: Some Surprises from Big Data”

Big Data and Business Analytics.
Session Chair: Per-Olof Arnäs, Chalmers

10:00 – Åke Ljungdahl, Volvo Car Group, Sweden. “Advanced Analytics at Volvo Cars”

10:30 – Coffee, sandwich, exhibition and posters

11:00 – Klas Eckervald, Teradata Corporation, Sweden. “Competitive Advantage by analyzing ALL your Data”

11:50 – Ib Groth-Clausen, Astra Zeneca, Sweden. “Disease Networks”

12:20 – Tobias Lehtipalo, Tibco SpotFire, Sweden. “The history of Spotfire - from dynamic queries to big data analytics for everyone”

12:50 – Lunch, exhibition and posters



Quality of Government and Human Well Being: Some Surprises from Big Data
Bo Rothstein

Issues about corruption and the quality of government where until recently neglected in the social sciences. This has changed dramatically thanks to the availability of “big data” which has produced a number of unexpected results that will be presented in this talk.

Bo Rothstein holds the August Röhss Chair in Political Science at University of Gothenburg where he is Head of the Quality of Government (QoG) Institute. The QoG Institute has compiled and made freely available one of the world's largest social science datasets that contain all relevant types of data for measuring QoG together with data measuring the effects of QoG for human well-being.



Advanced Analytics at Volvo Cars
Åke Ljungdahl

Volvo Cars is regarding IT and data as a strategic assets. We will show how this together with Advanced Analytics provides value in several business areas, and take a look at the challenges ahead.

Åke Ljungdahl is Vice President, Corporate Quality, and as such responsible for Customer Satisfaction Analytics, Quality Systems and Improvement Methodology within Volvo Car Group. He holds a Master Degree in Electrical and Electronics Engineering, and has some 30 years' experience from Product Development, Project Management, Quality Management and Manufacturing Engineering.



Competitive Advantage by analyzing ALL your Data
Klas Eckervald

Analyzing sensor networks, social media data, and other non-traditional data enables companies to gain more insight and potentially change the way they operate. Learn about a data-centric approach to leverage new algorithms and visualizations.

Klas Eckervald helps leading European organizations getting value out of Data by leveraging the latest advances in DW research. As a member of the Teradata Applied Intelligence & Innovation team, he covers Unified Data Architecture to enable Big Data Analytics, Cost Optimization via consolidation, Global DW challenges and High Performance Computing leveraging in-database capabilities for exceptional scalability.



Disease networks
Ib Groth-Clausen

The data explosion in Life Science continues. The talk will address the diversity and dimensionality of data and approaches to derive useful hypotheses of disease mechanisms.

Ib Groth Clausen has worked for 20 years with many aspects of Life Science Informatics. In his current role as Principal Scientist with Astra Zeneca he works with integrative approaches to uncover underlying mechanisms of inflammatory and autoimmune diseases to generate new targets for disease intervention and to discover new disease biomarkers.



The history of Spotfire – from dynamic queries to big data analytics for everyone
Tobias Lehtipalo

TIBCO Spotfire turns data into insights – Fast. Instantly visualize, interact with, and share data to spot the buried opportunities and risks nobody else can see. Bring insight to everyone by giving your data a voice that projects a clear message and encourages exploration. Anticipate what's next and quickly discover what you need to do to gain competitive advantage. Collaborate in context to accelerate more informed and transparent decisions.

Tobias Lehtipalo is Director of Product Management at TIBCO Software where he oversees product planning for the Spotfire product group. He graduated from Uppsala University with a Master of Science in Molecular Biotechnology Engineering and joined Spotfire as a developer in 1999. Over the years he has held various management positions in engineering, product strategy and marketing. He is passionate about user experience and about solving real problems for real people.



Illustration: Simon Fletcher

Programme

Big Data in Social Sciences and Digital Humanities. Session chair: Lars Borin, University of Gothenburg

14:00 – Stephan Meyer, IBM. “Inside Watson™”

14:30 – Antonella Fresa, Technical Coordinator of DCH-RP European Project, Italy
“Big Data in the Digital Cultural Heritage”

15:00 – Matthew Jockers, Department of English, University of Nebraska, USA
“Big Data and the Humanities”

15:30 – Coffee, pastries, exhibition and posters

16:00 – Panel Discussion. “Educating for the Big Data Era”. Moderator: Lena Peterson, Dept. of Computer Science and Engineering, Chalmers
– Daniel Langkilde, MSc student, Chalmers, Sweden
– Malte Isaksson, Volvo Car Group, Sweden
– Klas Eckervald, Teradata Corporation, Sweden
– Lars Borin, Centre for Language Technology, University of Gothenburg, Sweden
– Jan Smith, The network IT University, a collaboration between University of Gothenburg and Chalmers, Sweden
– Paweł Woźniak, PhD student, Chalmers, Sweden

17:00 – Closing remarks



Inside Watson™
Stephan Meyer

Watson™ is a computer system that facilitates a breakthrough in natural language processing. The process evolves from knowledge base acquisition and question analysis, towards finding the right answer out of hundreds of thousands of potential candidates.

Dr. Stephan Meyer is a Managing Strategy Consultant at IBM Global Business Services within the Advanced Analytics Center of Competency focused on driving speed and efficiency of business decisions through leveraging of advanced analytics practices such as modeling, predictive analytics, and optimization. Before joining IBM, Stephan worked as a research associate within the information management department at Karlsruhe Institute of Technology.



Big Data in the Digital Cultural Heritage
Antonella Fresa

Supply and demand of large datasets for academic and scientific research is becoming a constant reality in many scientific domains, also within the digital cultural heritage. The presentation aims to discuss which is in this scenario the role of museums, libraries and archives and how they are looking at a profitable cooperation with e-Infrastructures.

Antonella Fresa is Director at Promoter S.r.l., a small engineering company in Pisa (Italy). Since 2002, Technical Coordinator and Communication Manager of numerous European projects in the domain of digital cultural heritage, digital humanities, smart cities, digital preservation, e-Infrastructures. Previously, Antonella Fresa was Project Officer at the European Commission and Multimedia Development Manager at Tower Tech S.r.l. in Pisa.



Panel Discussion: Educating for the Big Data Era
Daniel Langkilde

Daniel Langkilde is a MSc student of Computational Science. He is curious on data mining, macroeconomics and putting information to good use. Daniel is a member of the board of Chalmers and Secretary of the Student council at Royal Swedish Academy of Engineering Sciences.



Panel Discussion: Educating for the Big Data Era
Klas Eckervald

Klas Eckervald (at Terra Data) helps leading European organizations getting value out of Data by leveraging the latest advances in DW research. As a member of the Teradata Applied Intelligence & Innovation team, he covers Unified Data Architecture to enable Big Data Analytics, Cost Optimization via consolidation, Global DW challenges and High Performance Computing leveraging in-database capabilities for exceptional scalability.



Big Data and the Humanities
Matthew Jockers

Jockers will present a humanistic perspective on how text mining, and what he calls "macroanalysis," can inform and enhance how we study literature.

Matthew L. Jockers is Assistant Professor of English at the University of Nebraska, Faculty Fellow in the Center for Digital Research in the Humanities, and Director of the Nebraska Literary Lab. He oversees UNL's post baccalaureate Certificate in Digital Humanities, and he serves as the faculty advisor for the minor in Digital Humanities. Prior to Nebraska, Jockers was a Lecturer and Academic Technology Specialist in the Department of English at Stanford where he co-founded the Stanford Literary Lab with Franco Moretti. Jockers's research is focused on computational approaches to the study of literature, especially large collections of literature.



Panel Discussion: Educating for the Big Data Era
Malte Isaksson

Malte Isaksson, MSc in Chemistry from University of Gothenburg. Malte Isaksson has worked with development and Advanced Engineering of light metals, corrosion and surface treatment. He is also Trained Six Sigma Black Belt, Six Sigma Master Black Belt and responsible for Six Sigma in Product Development Systems. From 2009 Corporate Master Black Belt/ Deployment Director and responsible for Six Sigma within VCC.



Panel Discussion: Educating for the Big Data Era
Lena Peterson

Lena Peterson is Dean of Education at Chalmers University of Technology. She is responsible for the education in the areas of Electrical and Computer Engineering and Software Engineering and Management and organization of technology. She is also associate professor at the Department of Computer Science and Engineering, Chalmers.



Panel Discussion: Educating for the Big Data Era
Lars Borin

Lars Borin is Professor of Natural Language Processing at the University of Gothenburg (GU), director of the Swedish Language Bank (offering online access for research to billions of words of linguistically processed texts representing all historical stages of Swedish), director of GU's Centre for Language Technology, and the national coordinator of the Swedish node of the ESFRI CLARIN ERIC.



Panel Discussion: Educating for the Big Data Era
Jan Smith

Jan Smith is Professor in Computing Science at Chalmers. Since 2006, he is responsible for IT University, a collaboration between Chalmers and University of Gothenburg, for applied information technology within education, research, society and business. He is also Dean at the IT faculty of University of Gothenburg. Through his assignments, Jan Smith has a vast knowledge of IT within academia in the West Sweden area.



Panel Discussion: Educating for the Big Data Era
Paweł Woźniak

Paweł Woźniak is a doctoral student in Human-Computer Interaction at Chalmers and part of the Data-Intensive Visual Analytics EU ITN Programme (DIVA). His research interests span a wide area from persuasive technology to novel mobile interaction techniques. He is particularly interested in enabling larger audiences to explore community-relevant data sets through interactive technologies.

**Main Partner and Sponsor:
Chalmers Area of Advance ICT**

**Further Partners and Sponsors:
Chalmers e-Science Centre
Chalmers Area of Advance Life Science Engineering
Chalmers Area of Advance Transport
Marie Curie Actions of FP7/2007-2013/, REA grant no 290227, DIVA
University of Gothenburg
Visualiseringscentrum Chalmers
Swedish Foundation for Strategic Research, IIS11**

**Thanks to:
Walter & Eliza Hall Institute of Medical Research, Australia
University of California, San Diego, USA
Region Västra Götaland, Sweden
GE Healthcare/Global Operations, Sweden
Microsoft Research, USA
Linköping University, Sweden
Recorded Future, Sweden
Oxford e-Research Centre, United Kingdom
Oxford Internet Institute, United Kingdom
Teradata Corporation, Sweden
Karlstad University, Sweden
MedTech West, Sweden
Astra Zeneca, Sweden
Volvo Car Group, Sweden
Tibco SpotFire, Sweden
IBM Global Business Services, Germany
DCH-RP European Project, Italy
University of Nebraska, USA**

**Organizing Team:
David Sands, Professor of Computer Science;
Co-director, Chalmers Area of Advance ICT
Devdatt Dubhashi, Professor of Computer Science,
Dept. of CSE
Morten Fjeld, chair Professor of
Human-Computer Interaction, Dept. of Applied IT
Karin Lundberg, Communications Officer,
Chalmers Area of Advance ICT
Martin Markström, Co-director,
Life Science Engineering Area of Advance
Karolina Partheen, Operations Officer,
Chalmers Area of Advance ICT**

**Poster Programme Chair: Paweł Wozniak,
PhD-student of Human-Computer Interaction,
Dept. of Applied IT**

**INFORMATION AND
COMMUNICATION TECHNOLOGY**
A CHALMERS
AREA OF ADVANCE

TRANSPORT
A CHALMERS
AREA OF ADVANCE



CHALMERS
e-Science Centre

**LIFE SCIENCE
ENGINEERING**
A CHALMERS
AREA OF ADVANCE



UNIVERSITY OF
GOTHENBURG

