

SDS|2017

4th Swiss Conference on Data Science Program

June 16th, 2017



Foreword

Welcome to SDS|2017, the 4th Swiss Conference on Data Science, organized by the Swiss Alliance for Data-Intensive Services (Data+Service).

Switzerland has a long tradition of creativity, innovation and entrepreneurship resulting in break-through scientific insights, world-class companies and products. It is the ambition of Data + Service and SDS|2017 to extend this success story by making Switzerland an internationally recognized hub for data-driven value creation. We at D ONE share this vision and feel strongly committed to the conference and Data + Service.

We wish you a rewarding conference, interesting encounters, exciting insights and great inspirations at SDS|2017. We hope it will have a positive impact on your personal and professional development as we are convinced: data-driven value creation is the key to success in the 21st century.

Dr. Simon Hefti Chairman & Co-Founder D ONE Hans Peter Gränicher CEO & Co-Founder D ONE

Contents

Talk information	4
Program	24
Sponsors & Supporters	30
Organizers	44
Conference parallel sessions	45
Notes	51
Imprint	52

Claudia Perlich Chief Scientist, Dstillery

Predictability and other Predicaments in Machine Learning Applications

Abstract.

In the context of building predictive models, predictability is usually considered a blessing. After all – that is the goal: build the model that has the highest predictive performance. The rise of 'big data' has in fact vastly improved our ability to predict human behavior thanks to the introduction of much more informative features. However, in practice things are more differentiated than that. For many applications, the relevant outcome is observed for very different reasons. In such mixed scenarios, the model will automatically gravitate to the one that is easiest to predict at the expense of the others. This even holds if the predictable scenario is by far less common or relevant. We present a number of applications where this happens: clicks on ads being performed 'intentionally' vs. 'accidentally', consumers visiting store locations vs. their phones pretending to be there, and finally customers filling out online forms vs. bots defrauding the advertising industry. In conclusion, the combination of different and highly informative features can have significantly negative impact on the usefulness of predictive modeling.

Biography.

Claudia Perlich leads the machine learning efforts that power Dstillery's digital intelligence for marketers and media companies. With more than 50 published scientific articles, she is a widely acclaimed expert on big data and machine learning applications, and an active speaker at data science and marketing conferences around the world.

Claudia is the past winner of the Advertising Research Foundation's (ARF) Grand Innovation Award and has been selected for Crain's New York's 40 Under 40 list, Wired Magazine's Smart List, and Fast Company's 100 Most Creative People.

Claudia holds multiple patents in machine learning. She has won many data mining competitions and awards at Knowledge Discovery and Data Mining (KDD) conferences, and served as the organization's General Chair in 2014.

Prior to joining Dstillery in 2010, Claudia worked at IBM's Watson Research Center, focusing on data analytics and machine learning. She holds a PhD in Information Systems from New York University (where she continues to teach at the Stern School of Business), and an MA in Computer Science from the University of Colorado.

Clemens Cap University of Rostock

Do We Require an Ethics of Data Science?

Abstract.

Data science methods appear as silver bullet for all kinds of empirical, data-centric research questions. As with the introduction of all new technologies, ethical and social implications should be studied. A closer examination of data science shows unpleasant risks and side

effects which reach far beyond mere privacy considerations.

The talk provides examples why some people are skeptical of the data science approach.

But as fear and doubt never are good answers to social and technological innovation, a path must be found where the important scientific and technological developments of data science can be used in a non-damaging way. This attempt leads, on the one hand, to philosophical and value centered questions and calls for an ethic of data science. However, it also requires additional research on the technological and product side. A successful application of data science must combine both aspects.

Biography.

Clemens H. Cap was born in Innsbruck, Austria. He studied mathematics, computer science and physics at the University of Innsbruck. He spent his postdoc years at the University of Zurich, Switzerland, where he acquired his habilitation (lecturer status) in the area of computer science. He served for one term as associate professor at the University of Mannheim, Germany. Since 1997 he is full professor for information and communication services at the University of Rostock, Germany and continues to be an adjunct professor at the University of Zurich. His research interests are applications of networked systems, security and social impacts of computing technology.

Philip van Hövell, Sarah Grimm PwC Switzerland

Track 1 11:00 - 11:45

Predictive Healthcare with Real-world Evidence for Neurological Disorders

Abstract.

A recent study led by the World Health Organization (WHO) pointed out that, primarily due to demographic shifts i.e. an ageing population, the burden of neurological disorders such as dementia and multiple sclerosis (MS) on society is expected to increase dramatically. Due to the high complexity and uncertainty of disease progression of neurological disorders, it currently takes more than one year to find an appropriate therapy for an MS patient. Our solution to support patients and their doctors during this difficult journey is to distill the large amounts of available data into meaningful and relevant decision-making information as efficiently as possible.

Our PHREND App contains a data-driven mathematical regression model to predict Multiple Sclerosis (MS) disorder progression for individuals based on different therapies. A network of doctors in Germany provides the data, while PwC develops and updates the model. Data for the model comes from 1) the individual patient's personal characteristics, and 2) Real-world evidence of therapy effectiveness for 20,000 MS sufferers.

The output of PHREND is an independent recommendation for the therapy that is statistically most likely to succeed for each individual patient, presented in a transparent and easy-to-understand way. We have validated our model by quantitative statistical analysis and provide evidence that the predictions of our model add value when accounting for personal characteristics on a patient-level.

We present a solution to support doctors when making treatment decisions for their MS patients. We are currently implementing the PHREND App in German doctors' offices, and plan to expand our model to several other neurological disorders.

Biography.

Philip van Hövell is a fully qualified actuary (SAV) (statistics), ETH Zurich and has an Msc in mathematical finance, University of Konstanz. He's been working at PwC Switzerland for several years and has lately developed a bio-statistical team within the Digital Services Business Unit. Primarily focusing on multiple sclerosis, Philip designed and managed the project on building our innovative tool to predict therapies for individual patients based on real world evidence (RWE) data. He also manages quant projects within the Pharma & Life Science industry.

Sarah Grimm joined PwC Switzerland's Data Analytics & Modelling team in February 2016 and has been working on predictive healthcare applications. Before joining PwC, she received her MSc in statistics from the Seminar für Statistik at ETH Zürich, where she afterwards also was a statistical consultant. During this time she gained broad experience in analytics through projects in a variety of fields using different methodologies to solve real world problems.

Olivier Heuberger-Götsch Scigility

Track 2 11:00 - 11:45

The New Swiss Data Protection Regulation in the digital age – A new data protection landscape?

Abstract.

The Federal Council presented on December 21, 2016 the proposal for a new Swiss data protection law, the Federal Act on Data Protection (E-FADP), which shall not come into force until the spring of 2019. However, the new EU data protection framework – the General Data Protection Regulation (GDPR) – will replace the current EU Directive and comes into force already on May 25, 2018. The GDPR will directly impact many Swiss companies, and thus the new Swiss data protection law will be adopting some of the main aspects of the GDPR. As the GDPR as well as the new Swiss data protection law contain a number of onerous obligations that will have a far-reaching impact, especially regarding a new chapter containing criminal charges, it will have an immediate impact for companies in the EU as well as in Switzerland.

The revision of the FADP, as well as the GDPR on a European level, will have a material and significant impact on how companies will process data in the future. This speech shall outline the emerging role of big data in our changing world and the opportunities and challenges that it presents in light of the forthcoming regulations. Now is the time to have a closer look at what we all will have to change.

Biography.

Olivier Heuberger-Götsch is a lawyer mainly active in the fields of ICT and intellectual property law, with focus on privacy and data protection law, licensing, internet and software law.

As Legal Counsel of Scigility, he advises companies, individuals and public-sector organizations in all fields of ICT/IP and Contract Law. He studied Law and Economics at the Zurich University of Applied Science and made his master's degree in Law at the University of Lucerne in 2009. He is admitted to the bar since 2012 and works since then as a lawyer and as Legal Counsel for Scigility. He is currently writing his PhD on IT-Law at the University of Lucerne.

Tom Bachmann, Christian Erni Smide, D ONE Solutions

Track 3 11:00 - 11:45

A Data Driven Business Model For E-Bike Sharing

Abstract.

Connected Mobility, Track and Trace, Internet of Things are common buzzwords. To properly understand this field of data-driven projects, Die Mobiliar has created a new business model for sharing E-Bikes that leverages bike position and user behaviour insights. Following a successful pilot involving its employees, it is now opening the scheme to consumers in Zurich.

In this talk, Tom Bachmann (Smide) and Christian Erni (D|ONE) explain how the business model was developed using the collected data during various pilot schemes.

Biography.

Christian Erni is a pioneer of new media communication and has a vast experience at the intersection of information and design. Christian joined D ONE as a member of the management team. His areas of expertise are data experience, marketing & communication.

Before his career with D ONE, Christian was Creative Director at international agencies and lecturer for visual design at Zurich University of the Arts. He won many international awards in design and was juror at different Festivals, eg. New York Festivals and presides the Best of Swiss Web jury "Creation".

Thomas Röfer University of Bremen

Track 4 11:00 - 11<u>:45</u>

RoboCup: Soccer as a Challenge for Artificial Intelligence,

Abstract.

From the early beginnings of computer science, researchers have tried to build systems that are more capable than humans in specific tasks. In particular, tasks have been addressed that are usually associated with intelligence. These have often been games. For instance, computer systems are now the best players in Chess, Jeopardy, Go, and Poker. However, decades of research in artificial intelligence have also revealed that many everyday tasks are much harder to master for artificial systems than tasks that are hard for humans. Therefore, the RoboCup Federation has set itself the goal that "by the mid-21st century, a team of fully autonomous humanoid soccer players shall win the soccer game against the winner of the most recent World Cup". This is very ambitious, because soccer players do not only need a mind, they also need a body, and they act as a team. The problem is addressed with several different soccer leagues, some just simulated, some with driving robots, and some with humanoids. The talk will give an insight in how such systems work with the 2016 Standard Platform League world champion team B-Human as an example.

Biography.

Thomas Röfer received his Diploma degree and PhD degree in computer science from the University of Bremen in 1993 and 1998. He was member of the DFG Transregional Collaborative Research Center "Spatial Cognition: Reasoning, Action, Interaction" between 2003 and 2014. Since 2006, he has been with the German Research Center for Artificial Intelligence (DFKI) in the department for Cyber-Physical Systems. He has been active in RoboCup since 2001, first as the speaker of the three-times world-champion "GermanTeam" in the Four-legged League, then as the leader of the five-times world champion team "B-Human" in the Standard Platform League. His research interests include humanoid robots, real-time computer vision, robot simulation, and behavior control.

Dimopoulos Sotiris

IT-Logix

Track 1 11:45 - 12:30

Real-Life Data Science Solutions for Mid-sized Swiss Companies

Abstract.

Nowadays, a great deal of small or mid-sized non-tech companies have embraced business intelligence solutions for data storage, reporting and visualization purposes. Yet, important daily operations or decision-making are not performed in a data-driven manner. but rather are based on experience and gut-feeling. Such companies often lack the resources to build the necessary infrastructure or to employ the personnel to perform data analytics, however, recent advances in the area of cloud computing together with the availability of modern fully automatic data analytics algorithms, allow practically any type of firm to easily leverage the power of data analytics. Here, we show real data science solutions applied to three small to mid-sized Swiss companies. First, we describe how a non-profit organization can apply marketing analytics in order to draw insight into the behavior of its donors, group them according to multiple dimensions and efficiently exploit its available marketing touchpoints. Second, we show how a retailer of photo-products can use data analytics to quantify the driving factors behind the sales, to assess the efficiency of the sales at the regional level and finally to use information on past transactions in order to build a recommendation engine for setting optimal discount values. Third, we describe a framework developed for a Swiss retailer of telecommunication products that aids its controlling department to process large batches of mobile contracts and detect automatically potentially fraudulent contracts. For all aforementioned cases, the whole operationalization process is also discussed by use of Microsoft Azure technology.

Biography.

Dr. Sotiris Dimopoulos currently works as a Data Scientist Consultant at IT-Logix AG, where he develops and designs data science solutions for various business partners. He holds a diploma on Electrical & Computer Engineering and in 2013 he was awarded a PhD from the department of Biosystems Science and Engineering, ETH Zurich. He has participated in numerous large-scale international research projects and is an author in various scientific publications. Since 2007 he lives in Zürich and is the founder and organizer of the Swiss Meet-up "Data Science for Business".

Jennifer Priestley Kennesaw State University, USA

Invited Talk Track 2

11:45 - 12:30

Corporate/University Partnerships: What Both Parties Need to Know

Abstract.

The private sector/university collaborative partnership is mutually beneficial and, in the current environment of resource-scare data science talent, is more important than ever – to both parties. In this talk, Dr. Priestley, will discuss the benefits (and potential pitfalls) of corporate/university partnerships and how both parties can get the most out of the collaboration. Dr. Priestley will provide case studies related to leveraging university programs at the undergraduate, masters and Ph.D. level.

Biography.

Jennifer Lewis Priestley is a Professor of Applied Statistics and Data Science at Kennesaw State University, where she is the Director of the Center for Statistics and Analytical Services. She oversees the Ph.D. Program in Advanced Analytics and Data Science, and teaches courses in Applied Statistics at the undergraduate, Masters and Ph.D. levels. In 2012, the SAS Institute recognized Dr. Priestley as the 2012 Distinguished Statistics Professor of the Year. She served as the 2012 and 2015 Co-Chair of the National Analytics Conference. Datanami recognized Dr. Priestley as one of the top 12 "Data Scientists to Watch in 2016."

Dr. Priestley has been a featured speaker at dozens of corporate events addressing issues related to advanced analytics and the challenges and opportunities of "Big Data". She has authored over 50 articles on Binary Classification, Risk Modeling, Sampling, Applications of Statistical Methodologies for Problem Solving as well as several textbook manuals for Excel, SAS, JMP and Minitab.

Prior to receiving a Ph.D. in Statistics, Dr. Priestley worked in the Financial Services industry for 11 years. Her positions included Vice President of Business Development for VISA EU in London, where she was responsible for developing the consumer credit markets for Irish and Scottish banks. She also worked for MasterCard International as a Vice President for Business Development, where she was responsible for banking relationships in the Southeastern US. She also held positions with AT&T Universal Card and with Andersen Consulting. Dr. Priestley received an MBA from The Pennsylvania State University, where she was president of the graduate student body, and a BS from Georgia Tech. She also received a certification from the ABA Bankcard School in Norman, OK, and a Certification in Base SAS Programming, and a Business Analyst Certification from the SAS Institute.

Gian-Marco Baschera

Value of Data Science – Ensure your precious time is wisely spent!

Abstract.

A data scientist's life can be extremely frustrating. Imagine, you are fully committed to a project and create fascinating predictive models that surpass all benchmarks, but the project gets stopped or the outcome is just never used. Why do we regularly observe this pattern?

Data science teams often operate too far away from the business and pursue use cases that have been developed based on the available data and technology, known as technology-push. We believe that data analytics use cases should be developed in a business-pull approach, ensuring that the data scientists work will ultimately create a business values.

In this talk we will focus on the very important – but often neglected – initial Vision & Scope phase of the Data Analytics Process. We will show approaches to identify use cases with high business value. We will discuss the different aspects that need to be considered and the deliverables that will be generated in this early phase of a data analytics project. Their importance will be illustrated based on lessons-learned from real-world data analytics projects.

Biography.

Gian-Marco Baschera is working as a lead data scientist at Zühlke and has several years of experience in conducting data analytics projects. Before joining Zühlke he worked as a data scientist at a brain & trauma foundation. Gian-Marco Baschera studied mathematics and computational science and engineering and holds a PhD in computer science in the field of statistical modelling and machine learning.

Jana Koehler Lucerne Univerity of Applied Sciences and Arts

Track 4 11:45 - 12:30

Intelligent Work Support for Industrial Service Desks with Cognitive Services

Abstract.

Service desks in industrial enterprises are challenged by an increasing complexity of skills required to effectively solve the problems of customers when industrial products are entering a connected world. Many companies consider outsourcing their service desk, although this means to lose the direct channel to the customer. In this talk, we argue that the knowledge of service desks is of tremendous value to enterprises and we show how artificial intelligence technologies can help service desks coping better with these challenges. We present 3 intelligent assistants that support service desk employees and we discuss to which extent currently commercially available cognitive services can help building these assistants.

Biography.

Jana Koehler is professor of computer science at Lucerne University of Applied Sciences and Arts, Switzerland. Prior affiliations include the German Research Center for Artificial Intelligence (DFKI), the Universities of Linköping (S), Maryland (USA) and Freiburg (D), Schindler Elevators R&D, and the IBM Research Laboratory in Zurich.

Her teaching and research work focuses on artificial intelligence, business process management, and software engineering and architecture. She currently conducts several industrial research projects with partners in the area of artificial intelligence. She is president of SGAICO, the Special Interest Group on Artificial Intelligence and Cognitive Science of the Swiss Informatics Society, was a contributor to the BPMN 2.0 and CMMN 1.0 standards, holds several patents and has an h-index of 37. Her publications are available via Google Scholar, DBLP or her own web page at http://user.enterpriselab.ch/~takoehle/

13:45 - 14:30

Rico Knapper Anacision

Successful implementation of a customized predictive maintenance solution in industrial manufacturing

Abstract.

Our customer, a large German engineering company with hundreds of plants worldwide, maintains its plants in fixed periodic intervals. Despite the daily collection of gigabytes of sensor data, our customer suffers significant economic damage due to unplanned and long shutdowns of plants.

In this project, our goal was to predict unexpected breakdowns of machine parts at early stages using real-time sensor data to maintain the plants just in time. This prevents damages of other parts and long unplanned shutdowns of the whole plant. However, the low number of total shutdowns in the past and the variety of the failures causing the shutdowns represented a major challenge in this project. Thus, applying common predictive maintenance approaches, which aim at identifying patterns

in the data that correlate with the historic incidents, was not feasible. Therefore, we developed a novel solution aiming at a continuous monitoring of the overall health state of the plants. For this purpose, we modeled physical dependencies between machine parts using regression models to predict physically expected sensor values. We than use the deviation between measured and predicted sensor data as input for a health function to quantify the health state of the plant. This enables us to predict plant failures in advance. Currently, we are implementing this novel predictive mai tenance approach to monitor the health state of industrial plants in real-time.

Biography.

Rico Knapper studied Mathematics at the Karlsruhe Institute of Technology (KIT), Germany. From 2009 till 2016 he worked as Research Assistant, Department Manager and Division Manager on applied Data Science projects at the FZI Research Center for Information Technology in Karlsruhe, Germany. In 2016, Rico Knapper co-founded the "anacision GmbH // an EXXETA company" and is currently involved in various applied Data Science projects in collaboration with different industrial customers.

Text and data mining for the development of improvement measures in industry

and services based on the Net Promotor Score Methodology

Abstract.

Manufacturers of high-quality products, need to keep up their reputation on the global stage while competition in world markets requires permanent product innovation, shorter product release cycles, reduced production cost and vigorous quality controls. The Net Promoter Score (NPS) is probably the easiest way to capture customer satisfaction as an early warning indicator. It's simple methodology was developed and published under the title "The one number you need to grow" by Frederick F. Reichheld from Bain & Company in the Harvard Business Review 2003. NPS-Analysis is based on data from a rolling survey with a very simple two question design:

- 1. Would you recommend our company / product / service to a colleague or friend? (rating scale: 0-10)
- 2. Please tell us why you gave that rating and let us know how we can improve it? With Text Analysis of answers to question 2 it is possible to discover new emerging issues and to weight their importance. Subsequent Prescriptive Modelling provides a root cause analysis estimating the TextTopic's impact on the numeric answer to question 1 which translates directly into the dynamics of customer recommendations and company growth. The final result consists in transparent indications for the most important improvements. Along this logical analysis chain we will explore in a case study the capabilities for the development of corrective actions that can be planned ahead of time on fact based decisions.

Biography.

Ulrich Reincke is a Principal Analytics Solution Architect in the Centre of Excellence at SAS Institute, Heidelberg. Over the last 18 years he has been responsible in different roles for analytical enterprise decision support solutions applying methods of statistics, machine learning, data mining, forecasting and optimization to empower organizations to perform better with fact-based decisions. Prior to SAS Mr. Reincke had positions in the Information Product Department of Deutsche Börse and in the Trade Division of The World Bank in Washington DC.

Volker Markl Technische Universitat Berlin

Invited Talk Track 3

13:45 - 14:30

Big Data Management and Apache Flink: Key Challenges and (Some) Solutions

Abstract.

The shortage of qualified data scientists is effectively limiting Big Data from fully realizing its potential to deliver insight and provide value for scientists, business analysts, and society as a whole. In order to remedy this situation, we believe that novel technologies that draw on the concepts of declarative languages, query optimization, automatic parallelization and hardware adaptation are necessary. In this talk, we will discuss several aspects of our research in this area, including results in how to optimize iterative data flow programs, optimistic fault-tolerance, and steps toward a deep language embedding of advanced data analysis programs. We will also discuss how our research activities have led to Apache Flink, an open-source big data analytics system, which by now has become a major data processing engine in the Apache Big Data Stack, used in a variety of applications by academia and industry.

Biography.

Volker Markl is a Full Professor and Chair of the Database Systems and Information Management (DIMA) group at the Technische Universitat Berlin (TU Berlin), director of the research group "Intelligent Analysis of Massive Data" at the German Research Center for Artificial Intelligence (DFKI), and speaker of the Berlin Big Data Center (BBDC). Earlier in his career, Dr. Markl lead a research group at FORWISS, the Bavarian Research Center for Knowledge-based Systems in Munich, Germany, and was a Research Staff member & Project Leader at the IBM Almaden Research Center in San Jose, California, USA. Dr. Markl has published numerous research papers on indexing, query optimization, lightweight information integration, and scalable data processing. He holds 19 patents, has transferred technology into several commercial products, and advises several companies and startups. He has been speaker and principal investigator of the Stratosphere research project that resulted in the "Apache Flink" big data analytics system. Dr. Markl currently serves as the secretary of the VLDB Endowment and was elected as one of Germany's leading "digital minds" (Digitale Köpfe) by the German Informatics Society (GI). Volker Markl and his team earned an ACM SIGMOD Research Highlight Award 2016 for their work on implicit parallelism through deep language embedding.

13:45 - 14:30

Artificial Intelligence for Accelerating and Improving Diagnostic Radiology

Abstract.

The diagnosis and treatment of cancer has been drastically improved by newer imaging methods like PET-CT which generate large number of images where single spots can drastically influence the diagnosis and treatment.

For physicians this means a long time must be spent carefully reading thousands images a day and looking at dozens of different regions carefully to check for the possibility of aggressive disease.

4Quant Ltd. together with the University Hospital Basel has demonstrated the potential to radically reduce the physicians reading time without sacrificing quality by using a Deep Learning approach. We present the work we have done towards computer aided staging of Non-Small Cell Lung Cancer (NSCLC) for a more efficient and evidence based precision medicine and illustrate challenges, hurdles and findings while developing an Al-based product for medical use.

Biography.

Kevin Mader studied Electrical Engineering and Photonics at Boston University and did his PhD in Biomedical Engineering at ETH Zurich. He won the Pioneer Fellowship for entrepreneurship from ETH Zurich and subsequently founded 4Quant, a company that does Big image analytics. Kevin is also a lecturer in the X-ray Microscopy Group within the Biomedical Engineering Institute at ETH Zurich. His research focuses on turning big hairy 3D images into simple, robust, reproducible numbers. In particular, as part of several partnerships, he is currently working on automatically segmenting full animal zebrafish images, characterizing rheology in 3D flows, and measuring viral infection dynamics in cell lines.

Flavio Trolese studied Computer Science at the University of Applied Science in Zurich, cofounded and built up two software companies and joined Kevin in 2015 supporting him to elaborate a marketable technical platform to process high volumes of image and video data based on machine learning and to work on sales and business development. 4Quant Ltd. exists since 2015 as a spin-off of ETH Zurich and PSI.

Mariam Barque

University of Applied Sciences Western Switzerland

14:30 - 15:15

Short-Term Renewable Energy Prediction using Machine Learning Algorithms

Abstract.

Photovoltaics (PV) and other renewable energy sources have widely spread in recent years. Although those sources provide an environmentally-friendly solution, their integration in the electric grid is a real challenge as they depend on meteorological conditions. Therefore, the ability to predict those variable sources considering meteorological uncertainty plays a key role in the management of the energy supply needs and reserves. Our research lab aims at providing innovative solutions for renewables integration. We target user-oriented methodologies from the collection of data to the prediction implementation and visualization. We have performed solar and wind production prediction tools and have been able to implement them for ESR (Energies Sion Region) and SEIC-Teledis, which are regional electricity providers. We would like to share our data intensive solution to predict renewable energy using time series analyses and Machine Learning Algorithms.

Our goal is to provide a forecasting model to set the day-ahead grid electricity needs. This information plays a key role in power dispatching plans and grid charge control. The main novelty of our approach is to provide a user-friendly and flexible solution that combines classification algorithms and a nonlinear regression. Hourly historical weather measures and production data have been used as training inputs whereas weather forecast parameters have been used as prediction inputs. To improve weather prediction results, past weather forecast errors have been trained and predicted using a nonlinear regression. Several algorithms like ANN, Random Forest, Gradient Boosting trees and linear regression where applied on the data of the 205 kW PV production of the microgrid in the Techno-Pôle of Sierre. Considering wind power prediction, our test and application data are the production of the three wind chills in Valais managed by SEIC-Teledis, the regional electricity provider, from whom 4 years of hourly wind production data were collected. As a result of the work, a user-friendly visualization which is daily updated with the production forecast have been deployed for the grid operator.

Biography.

Mariam Barque is currently Research Assistant at the Information System Institute of HES-SO in Sierre and works on energy management related projects. Her work aims at analyzing and forecasting energy production and consumption for microgrids management or grid-operators support. Previously consultant in Computerised Maintenance Management Systems (CMMS) training and deployment for EDF (Electricité de France), she also modelized the Carbon Capture and Storage technologies in EDF's long term energy model in order to analyse the impacts of such technologies in energy transition by 2050 in terms of electricity costs and mix. With a strong interest in energy systems and decision-making related tools, Mariam is an energy engineer of Ecole des Mines d'Albi-Carmaux with a specialized mastere degree in Energy Systems Optimization of Mines ParisTech.

Stijn Vermeeren, Iwan Gloor Starmind International

Track 2 14:30 - 15:15

Starmind: Using artificial intelligence to leverage human ingenuity

Abstract.

There have been huge advances in artificial intelligence in recent years, yet solving the majority of problems still requires unique human skills. Finding the right people with the right skills for a problem at hand can be difficult, especially within large organisations. However, even when machine learning cannot directly solve the problem, it can still help with finding the right expert who can provide the solution. Starmind is a successful Swiss startup which uses machine learning to automatically find the right experts on any given topic. The Starmind technology is currently in use in over 40 countries by renowned companies from various industries such as Swisscom, J. Walter Thompson, Bayer etc. This talk will present how Starmind combines existing natural language processing techniques with specialized algorithms to create a self-learning expertise platform. It will also outline Starmind's vision of how such technologies will become increasingly integrated in the working lives of a growing number of employees, thereby improving their productivity.

Biography.

Stijn Vermeeren has been working as Lead Algorithm Engineer for Starmind since 2013. Previously, he studied mathematics at the universities of Leuven and Cambridge before obtaining a PhD in mathematical logic from the University of Leeds.

Iwan Gloor is the Lead Solution Architect at Starmind. His background is in theoretical physics, web development and artificial intelligence. Before Starmind, he worked for a global technology consulting company.

Jeremy Caliner SIX Group

SIX Market Data Detective: automated price anomaly detection for market data quality monitoring

Abstract.

SIX Financial Information delivers on an average day about 8.5bn price messages or over 600GB of financial market data to its customers through its market data feed. The data is collected and aggregated from over 1400 stock exchanges, multilateral trade facilities and over the counter markets.

Data quality and speed are central to SIX's value proposition. SIX currently has a simple price movement filter combined with a team of data quality specialists performing manual validation of the data. The Market Data Detective is an initiative to use modern data science tools to validate the market data before it enters the feed. The goals of the Market Data Detective are to increase the quality of data delivered to customers, increase the transparency of data quality, and reduce the manual workload and costs. Here we discuss the challenges and potential algorithmic solutions.

Biography.

Jeremy Callner, with over 10 years' experience in data science, physics and finance, is leading the effort at SIX Group to generate value from its large data sets. SIX Group owns and operates the Swiss Exchange, Swiss Clearing and Settlement Systems, Payment Systems in Switzerland, Austria, and Luxembourg, and also serves as a worldwide financial data provider. Before joining SIX in 2014, Jeremy spearheaded the methodology and technology for estimating and simulating from large covariance matrices of financial time series at swissQuant AG in Zurich. He also contributed to the development of some of swissQuant's most successful commercial Risk Tools.

Jeremy's previous work at the Large Hadron Collider at CERN with the CMS Collaboration focused on measuring the correlations of particle trajectories in heavy ion collisions in order to better understand the state of matter in the microseconds after the big bang. Jeremy also designed and tested many of the algorithms used to filter and clean the collision data both online and offline. Jeremy holds a PhD in Physics from the University of Illinois at Chicago and a bachelor's degree in Jazz Saxophone from Roosevelt University.

Thilo Stadelmann, Oliver Dürr

Track 4 14:30 - 15:15

Zurich University of Applied Sciences

Beyond ImageNet - Deep learning in industrial practice

Abstract.

Deep Learning (DL) methods have gained considerable attention in the Data Science community in the recent 5 years. In this talk are briefly reviewing the state of the art in DL and then give several examples of applications from diverse areas of industrial practice. We will focus on convolutional neural networks (CNNs), which have since the seminal work of Krizhevsky et al. in 2012 revolutionized image classification and are now surpassing human performance on some benchmark data sets. We will first showcase more classical applications of CNNs on image-like data, in particular phenotype classification of cells based on their morphology. While deep neural networks have become popular primarily on image classification tasks, they can also be successfully applied to other areas and tasks where there exists some local structure in the data. This includes for example outlier detection in a predictive maintenance setting, and semantic segmentation. We will highlight an application to segment newspaper pages into its constituting articles based on clues in the pixels, before we conclude by giving advice on how to work with DL having limited resources (e.g., training data).

Biography.

Thilo Stadelmann is a senior lecturer of computer science at ZHAW School of Engineering in Winterthur. He received his doctor of science degree from Marburg University in 2010, where he worked on multimedia analysis and voice recognition. Thilo joined the automotive industry for 3 years as a software architect, consultant and part of the management team of a 200-employee Swabian SME prior to switching back to academia. His current research focuses on applications of machine learning to diverse kinds of data. He is head of the ZHAW Datalab and vice president of SGAICO, the Swiss Group for Artificial Intelligence and Cognitive Sciences. See www.zhaw.ch/=stdm

Oliver Dürr is a senior lecturer of statistical data analysis at the ZHAW School of Engineering in Winterthur. In his applied research at ZHAW he focuses on deep learning. He lectures applied statistics and machine learning for industrial engineers (Studiengang Wirtschaftsingenieurwesen) and deep learning for the CAS Machine Intelligence. Before that he worked 9 years for Genedata (a swiss bioinformatics company) as a senior algorithm developer and consultant for statistics and machine learning, where his area of work was mainly gene expression analysis and screening. Oliver Dürr received his PhD from the University of Konstanz in 2003, where he researched in statistical / computational physics. See www.zhaw.ch/=dueo.

Matthias Mohler SAP

Applying predictive analytics in business: Our journey, customer examples and an outlook on self-service and automation

Abstract.

Initially coming from the ERP software space, SAP has become a strong player in the Swisscom is one of the largest integrators for SAP business software in Switzerland. We start our journey in 2015. A group of non-scientist recognizes the huge potential in the business process data of its' client systems. First proof-of-concepts are being conducted together with clients and with the help of experienced data scientist. Interesting projects follow with often surprising facts, interesting learnings and impressive business cases. We conclude with an overview of areas of the business world we believe can be improved with predictive algorithms.

In the second part of the presentation we discuss the benefits and risks for automating model creation: Automating data preparation and model selection give non-scientist a powerful tool in their hands. How can this be used to the best in real-life? What are the challenges when we have hundreds or thousands of predictive models in our company?

The tool used to make the predictions uses automation and hides a lot of the complexity. Matthias Mohler outlines upsides and risks of this. In the last part we dive into the possibilities of automatically deploying and maintaining models to increase precision and agility as well as keep the efforts low.

Biography.

As Head of Analytics Consulting at Swisscom Digital Enterprise Solutions, Matthias Mohler leads a team of professionals serving the Swiss market with expertise in SAP Business Objects and Predictive Analytics.

Prior to his current role, Matthias was advising corporate clients in business intelligence and data warehousing on behalf of IBM Global Business Services and Swisscom IT Services. As a senior consultant, he took over various project roles such as team lead, architect or developer. In addition to domestic assignments, Matthias was managing a global delivery project team based in The Netherlands. During his career he obtained extensive industry knowledge in real estate, oil & gas, industrial, IT services and telecommunications. Before moving into consulting, Matthias was working as a data analyst for the marketing department of Swisscom's telco division.

Matthias holds a professional college degree in business information systems as well as several internationally recognized professional certifications in enterprise architecture, project management, IT service management, business intelligence and SAP.

Ralph Badinelli

Pamplin College of Business of Virginia Tech

Invited Talk Track 2

15:15 - 16:00

The human-technology frontier for next-generation, smart service systems

Abstract.

What are the challenges on the horizon as the service economy, fueled by new technologies, transforms and disrupts the lives of people in every part of the world? This question motivates audacious research-and-development initiatives across industry, government and research institutions. My own research and involvement in professional societies has brought these initiatives into high relief.

In this presentation, I summarize the results of a recent Workshop sponsored by the U.S. National Science Foundation (NSF) and the International Society of Service Innovation Professionals (ISSIP), which was held to "Focus attention on research/technologies that neither industry nor academia are working on at the moment, but that may be necessary to enable the next generation of service systems." The outcome of this international gathering of thought leaders is an agenda for the next generation of industry and academic research in service innovations.

Two salient themes drive these initiatives. First, is the burgeoning awareness and understanding of the true nature of service as cocreation of value, which lies at the heart of every value-generating activity. Second, is the rapid and dramatic advent of new technologies, which enable the deployment of smart service systems. These themes enlighten some common errors in the perspectives behind many service innovation projects: connectivity alone cannot induce intelligence; technology enables, but does not create service; predictive models do not constitute decision support systems. We are consequently driven to model the human-technology symbiosis as the foundation of service innovation.

Biography.

Ralph Badinelli holds the Lenz Chair in the Department of Business Information Technology, Pamplin College of Business of Virginia Tech. His research interests are in the field of service science. He has published refereed articles in Operations Research, Service Science, Management Science, Decision Sciences, European Journal of Operational Research, Naval Research Logistics, Computers and Operations Research, International Journal of Production Research and other international journals. He is a member of the Institute for Operations Research and Management Science (INFORMS), Project Management Institute, former Chairperson of the INFORMS Service Science Section, Founding Board Member and President (2017) of the International Society for Service Innovation Professionals (ISSIP). His teaching responsibilities are in the areas of analytics, operations management and project management at the undergraduate, MBA and doctoral levels. He received a B.S. in mathematics and physics from Hofstra University, an M.S. in physics, an M.S. in business and a Ph.D. in business from Purdue University.

Program

	Time	Track 1	Track 2
	08:15	Welcome reception	
			Unified Tr
	08:45		Welc Swiss Alliance fo
L	09:00		Keynote: "Predictability and other Predicame
	09:45		Keynote: "Do We Require an I
	10:30	Short Break (Foyer Sopra)	
		Track 1 (Orione)	Track 2 (Bellavista 5)
	11:00	"Predictive Healthcare with Real-world Evidence for Neurological Disorders", Philip van Hövell, Sarah Grimm	"The New Swiss Data Protection Regulation in the digital age – A new data protection landscape?", Olivier Heuberger-Götsch
	11:45	"Real-Life Data Science Solutions for Mid-sized Swiss Companies", Dimopoulos Sotiris	Invited talk: "Corporate/University Partnerships: What Both Parties Need to Know", Jennifer Priestley
	12:30	Lunch Break (Foyer Sopra)	
	13:45	"Successful implementation of a customized predictive maintenance solution in industrial manufacturing", Rico Knapper	"Text and data mining for the development of improvement measures in industry and services based on the Net Promotor Score Methodology", Ulrich Reincke
	14:30	"Short-Term Renewable Energy Prediction using Machine Learning Algorithms", Mariam Barque	"Starmind: Using artificial intelligence to leverage human ingenuity", Stijn Vermeeren, Iwan Gloor
	15:15	"Applying predictive analytics in business: Our journey, customer examples and an outlook on self-service and automation", Matthias Mohler	Invited talk: "The human-technology frontier for next-generation, smart service systems", Ralph Badinelli
	16:00	Coffee Break (Foyer Sopra)	
ľ			Unified Tr
	16:30		Keynote: "Personal information system
	17:15		Keynote: "Spie
	18:00		Clos
	18:15	Apéro & Open End (Foyer Sopra)	

Track 3	Track 4			
ack (Szenario 1+2)				
oming address or Data-Intensive Services				
nts in Machine Learning Applications", Claudia Perlich				
Ethics of Data Science?", Clemens Cap				
Track 3 (Szenario 1)	Track 4 - AI in Industry (Szenario 2)			
"A Data Driven Business Model For E-Bike Sharing", Tom Bachmann, Christian Erni	Invited talk: "RoboCup: Soccer as a Challenge for Artificial Intelligence", Thomas Röfer			
"Value of Data Science – Ensure your precious time is wisely spent!", Gian-Marco Baschera	"Intelligent Work Support for Industrial Service Desks with Cognitive Services", Jana Koehler			
Invited talk: "Big Data Management and Apache Flink: Key Challenges and (Some) Solutions", Volker Markl	"Artificial Intelligence for Accelerating and Improving Diagnostic Radiology", Kevin Mader, Flavio Trolese			
"SIX Market Data Detective: automated price anomaly detection for market data quality monitoring", Jeremy Callner	"Beyond ImageNet - Deep learning in industrial practice", Thilo Stadelmann, Oliver Dürr			
"Automatic pre-analysis, metadata summary and visualisation for statistical mass data col- lections of unknown structure", Pierre Mandrin	"Machine Learning in the Interpretation of Meniscal Tear in Knee MRI", Chiara Civardi			
ack (Szenario 1+2)				
s and personal semantics", Gregory Grefenstette				
gelMining", David Kriesel				
ing and Apéro				

Pierre Mandrin

IMSD

Track 3 15:15 - 16:00

Automatic pre-analysis, metadata summary and visualisation for statistical mass data collections of unknown structure

Abstract.

We present an automatic pre-analysis and visualisation procedure for data files of unknown structure, which has been recently developed and tested. At the beginning of mass data processing with unknown structure, two important tasks must be performed: 1. the identification of the data structure (mostly as tables or lists) and 2. the assignment of the data to categories in order to allow for category-specific processing. The level of refinement of data processing crucially depends on how much data features (e.g. the category of statistical variables, the dependence between variables, the naming of variables or the frequency of occurrence of character patterns) can be associated to the data sequences. Although the association of data features to categories is similar to source-to-target schema mapping, no specific target objects are supplied, and thus the pre-analysis can only be informative by means of sophisticated categorisation tools. We demonstrate our procedure for the data sample of an experimental study on similarity based learning and reprocess the data using an automatic data preprocessor prototype which we have developed.

Biography.

Pierre A. Mandrin (Ph.D. in 1999 in physics) is a senior data scientist at IMSD (IMSD is a young service provider active in data analytics, statistics, Smart Data and predictive maintenance). P. Mandrin has initiated, conducted and published numerous projects in the fields of experimental, numerical and theoretical physics as well as didactics, one of which has been included in a metaanalysis. In his current projects at IMSD, he is developing tools for automatic reports of medical surveys and automated data pre-analysis and visualisation.

15:15 - 16:00

Machine Learning in the Interpretation of Meniscal Tear in Knee MRI

Abstract.

Medical imaging plays a key role in the diagnosis and treatment of diseases and health conditions, and technology breakthroughs have allowed advanced medical imaging, such as CT and MRI, to reach unprecedented resolutions. The number of devices and analyses increases every year. Nevertheless, this higher demand for medical imaging is not accompanied by a proportional increase in the number of radiologists, thus causing pressure, loss of accuracy, and the outsourcing of radiology. All factors that can possibly lead to a misdiagnosis.

Big datasets available from medical imaging provide the perfect background to test the potential of Artificial Intelligence and its recent advances. By feeding the machine with several medical cases as training data, and running supervised/semi-supervised/unsupervised learning algorithms, it is possible for the software to come up with classification rules based on connections it found within the data. Therefore, machine learning has the potential to improve radiology workflow efficiency, reducing costs and time allocated to medical imaging, and lowering the risk of diagnostic errors.

Through artificial intelligence, built into its product ScanDiags, Balzano helps reducing the cost and increasing the benefit of the interpretation of medical images.

ScanDiags assists radiologists in analyzing orthopedic MRI scans, reducing the time required for the typical diagnostic process. The product lists structures that can be detected in MRI sequences. It produces the raw structure of a report that can then be completed with further insights and a diagnosis by the radiologist or doctor.

ScanDiags also processes feedback that is submitted from finalized medical reports - and continues learning. An instance of ScanDiags represents the central repository of experience and methodology of any organization. It allows time savings, knowledge transfer, second opinion, retrospective quality-review, and further scenarios. Plus, ScanDiags never tires and works overtime at no additional cost.

Here we demonstrate the effectiveness of convolutional deep neural networks and supervised machine learning algorithms in musculoskeletal imaging, for the classification of meniscus tears in knee MRIs. Since deep learning does not depend on the type of image data used, it could be adapted to the detection of other diseases and health conditions, as well as to other imaging analyses, like CT, and other specialties.

Biography.

Chiara works as Communication Officer for Science & Technology at Balzano Informatics, company that applies Machine Learning to different businesses. In her previous research, Chiara applied Network Theory and Deep Neural Network to Archeology and Biology, and she has extensive experience in X-ray Computed Tomography and visualization techniques. Chiara holds a PhD from ETH Zurich. She received her BSc from the University of Parma, Italy, and her MSc from the University of Southampton, UK.

Gregory Grefenstette

Florida Institute for Human & Machine Cognition

Keynote 16:30 - 17:15

Personal information systems and personal semantics

Abstract.

Wearable device are capturing more of our everyday life, allowing us to passively create a "lifelog" record of our activities. Current research in lifelog data has not paid enough attention to analysis of cognitive activities in comparison to physical activities. We argue that as we look into the future, wearable devices are going to be cheaper and more prevalent and textual data will play a more significant role. Data captured by lifelogging devices will increasingly include speech and text, potentially useful in analysis of intellectual activities. Analysing what a person hears, reads, and sees, we should be able to measure the extent of cognitive activity devoted to a certain topic or subject by a learner. Test-based lifelog records can benefit from semantic analysis tools developed for natural language processing. We show how semantic analysis of such text data can be achieved through the use of taxonomic subject facets and how these facets might be useful in quantifying cognitive activity devoted to various topics in a person's day. We are currently developing a method to automatically create taxonomic topic vocabularies that can be applied to this detection of intellectual activity. We will illustrate this in a private personal information platform we have developed.

Biography.

Gregory Grefenstette is senior associate researcher at the Florida Institute for Human & Machine Cognition (IHMC). He is a leading researcher in the field of natural language processing, helping to pioneer the fields of cross-language information retrieval and of distributional semantics, the induction and extraction of meaning from large quantities of text. Previous to his association with IHMC, Dr. Grefenstette was Advanced Researcher at Inria, the French national research institute in Computer Science, working on personal semantics, automatically creating taxonomies of personal interest such as hobbies and illnesses, used to annotate big personal data generated by a person's interaction with the digital world. Prior to that he was Chief Science Officer of Exalead, a search engine company, managing the OSEO QUAERO CMSE program on innovative multimedia indexing. Former chief scientist at the Xerox Research Centre Europe (1993-2001), at Clairvoyance Corporation (2001-2004), and with the French CEA (2004-2008), Dr. Grefenstette has been active in transferring research into products, and is named as inventor in 20 U.S patents. His research teams have been awarded the ACM Multimedia Grand Challenge 2010 Bronze award for "Introducing topic segmentation and segmented-based browsing tools into a content based video retrieval system" the ACM Multimedia Grand Challenge 2009 Award Most Practical System: "VoxaleadNews: Robust Automatic Segmentation of Video into Browsable Content" and a Lagardere Foundation 2007 three year grant for "Semantic Maps". Dr. Grefenstette is the author and editor of the following books: Search Based Applications (with Laura Wilber, Morgan Claypool, 2011), Text- and Speech-Triggered Information Access (Ed. With Steve Renals, Springer 2003) Cross-Language Information Retrieval (Ed. Kluwer, 1998) Explorations in Automatic Thesaurus Discovery (Kluwer, 1994)

28

David Kriesel Procter & Gamble

SpiegelMining

Abstract.

Since 2014, David downloaded roughly 100.000 articles from SpiegelOnline - the largest news portal and maybe most powerful opinion former in the German-speaking world. Using this large data set, he provides funny, entertaining, surprising, scary and even slightly politically incorrect insight in both SpiegelOnline's behavior and Data Science itself - in a colorful way that everybody can understand.

Biography.

David Kriesel studied Computer Science in Bonn, Germany, and at Cornell University, Ithaca, New York. He has a passion for biology and social insects. In real life, he works as an Advanced Analytics Technology Engineer at Procter & Gamble and as a self-employed public speaker.

In computer science, especially the biology inspired topics like neurocomputation, machine learning and swarm behavior make his day; He also wrote a 250p Book about the one of the most sophisticated machine learning technique today, Neural Networks, which he made available for free at his home page in German and English language (look for "A Brief Introduction To Neural Networks").

Aside from computer science, David is very interested in how the digitalization of the World changes society. David gained quite a bit of international recognition by bringing the "Xerox Scanning Bug" to public attention, followed by consequently forcing its correction.

DONE Solutions

DATA DRIVEN VALUE CREATION



BUSINESS SCIENCE: THE PROCESS OF GAINING INSIGHTS FROM DATA





Just like chocolate, business science is all about carefully selecting your raw materials. But how do you do that with data?

The 6C framework is tried and tested:
Clean – Complete – Current – Consistent –
Collaborative – Compliant

Presenting Partner



Success lies not in technology, but in well-conceived and consistent data processing.

What data needs to be processed? How are the important elements defined? What calculations are required?



A Kingdom for a Big Data Idea is a tried-andtested framework. It suggests three possible ways to answer these questions:

- 1. Through processes
- 2. Through pain points
- 3. Through interactive, data-based brainstorming sessions



Appetite grows when food looks nice. And it is exactly the same with data.

Knowledge only works when it is understood. And it is only understood when there is clarity and transparency.

When we see a beautiful praline in front of us we know just what to do with it! Likewise, when we possess clearly-defined knowledge, we know exactly what action to take.



Just as with chocolate, business science offers a variety of pathways to success. 80% of managers say that data is important, but most don't know how to implement these kinds of projects successfully.









Das Framework Value & Impact - Organisation & Culture - Data - Technology hat sich bewährt.

Of course, it puts data front and centre. Quality is important and this can only be guaranteed with processes.

Stumbling blocks

- No clear idea of what value needs to be derived from the data.
- Failure to understand that both workplace culture and tasks change through working with data. Change provokes fear and resistance, both of which need to be managed.

In the end, you need technology. High-powered technology.

The Key Success Factors for Business Science on one page. Experience at work:





Data are the currency of the digital economy – capital just waiting to be put to work. Smart organisations are realising that they could be putting this capital to much more profitable use. At PwC, data are at the heart of everything we do. We have a wealth of experience in all data-related disciplines from collection, cleansing and management to building analytical algorithms and visualisation tools. We draw

on this expertise to help you harness the real power of data lies by creating actionable insight: information you can act upon to transform your organisation.

How do we approach data and analytics?

At PwC, we believe that there are three key components to effective data and analytics:

- Innovation Lab: helping organisations accelerate their data analytics innovation
- Business Solutions: helping organisations solve business problems with our data analytics apps
- **Strategy through Execution**: helping organisations develop and implement the right data analytics strategy

All three components of our approach are underpinned by PwC's own industry expertise and experience. We develop, try, test and implement various business solutions for our clients and for PwC, as well as delivering end-to-end data analytics transformation, from developing the right strategy through to execution. In Switzerland we're PwC's centre of excellence for machine learning, artificial intelligence, optimisation, simulation and modelling. We have alliances with Google, the Fraunhofer Institute and other universities, and also work with a number of strategic delivery partners. We offer complex descriptive, predictive and prescriptive analytics, and build apps used by PwC territories globally. And we leverage secure, scalable data analytics managed services from our solution centres across the world.

Reach out to our team of analytics experts:

Christian Westermann, Data & Analytics, +41 58 792 27 97 christian.westermann@ch.pwc.com

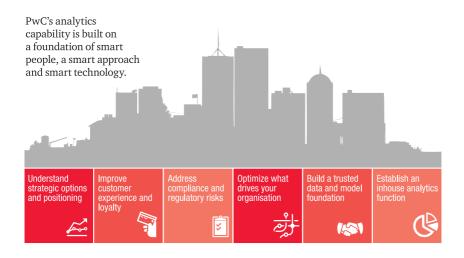
Jörg Gerigk, Data & Analytics, +41 58 792 27 19 joerg.gerigk@ch.pwc.com

Christian Spindler, Data & Analytics, +41 58 792 23 11 christian.spindler@ch.pwc.com

Platinum Sponsor

At PwC, our purpose is to build trust in society and solve important problems. We're a network of firms in 157 countries with more than 223,000 people who are committed to delivering quality in assurance, advisory and tax services. Within PwC Switzerland more than 3,000 employees and partners in 14 locations in Switzerland and one in the Principality of Liechtenstein help to create the value organisations and individuals are looking for.

PwC Digital Services in Switzerland is part of the worldwide community of experts dedicated to helping you face digital disruption and also to thrive on it. Our integrated digital solutions span innovation and strategy through to execution and include trust at every step along the way. We combine multi-disciplinary capabilities in digital strategy, transformation, user experience and design, cybersecurity as well as advanced analytics to help clients with all aspects of their digital transformation.





IT-Logix is your trustworthy Swiss Business Intelligence (BI), Data Warehouse and Big Data expert. We are recognized by SAP for Business Intelligence and by Microsoft (Gold Data Analytics) in Switzerland.

Our mission is to transform your structured and unstructured data into valuable business information. Your business users will be able to access, analyze and share information to make the decision process faster and more reliable each and every day. We know how to handle big volume, high variety and changing velocity to get maximum value from Big Data. In addition to requirements analysis and technical implementation, we will support you in visualizing the data according to the IBCS standards.

Our consulting services in BI strategy development, Business Intelligence Competence Center (BICC) organization set-up and agile project management complete our high quality end-to-end Business Intelligence services portfolio.

We are curious. We love to solve riddles. We challenge the status quo. We have common sense. We use proven methods and cutting-edge technologies to drive your business success.

This is why: We let your data speak.





Since its establishment in 1989, Microsoft Switzerland has developed from a small operation with three staff members into a mid-sized company with 620 employees, thanks to controlled and continuous growth. The headquarters of Microsoft Switzerland are located in Wallisellen near Zurich, additional

offices are located in Wollishofen, Bern and Geneva. The executive management is currently represented by Marianne Janik, Country Manager of Microsoft Switzerland. In global sales rankings, the Swiss subsidiary occupies 13th place out of 119 Microsoft subsidiaries. In the ranking of "Sales per PC", it even holds first place. This is first and foremost due to the conviction of many Swiss companies that the intelligent use of information and communication technologies leads to greater efficiency and competitive advantages.

Microsoft Switzerland has grappled with the changes in the working environment and its own work culture for years. As a result, in 2011 the headquarters of Microsoft Switzerland in Wallisellen were renovated on the basis of the latest findings. For the duration of the three-month renovation, Microsoft was involved in a special experiment: the office was completely closed and the employees shifted all of their activity to their home office or a location of their choosing. The result was not surprising, but worth noting: business activity was not negatively affected at all, but the social environment was. It was made clear that the office will continue to be of great significance in the future. Not as an individual workspace, but as a place for meeting and a platform for social interaction. The journey into the new working world is by no means finished – it has only just begun.

We are convinced that the centre of knowledge in Switzerland can only ensure and expand upon its success for the long-term if more modern technologies are used efficiently and intelligently to enable companies of all sizes to shape their very own "digital transformation". Microsoft Switzerland and its partner network are working on enabling and supporting Swiss entities to embrace their digital future.



What we do

We believe innovation requires an entrepreneurial mindset combined with the courage to push boundaries and break new ground.

Zühlke helps its clients to transform their vision from a smart idea into a resounding market success. Our tailor-made machine learning and analytics solutions support our customers in optimizing

their processes, help to make better, da-ta-based decisions and provide the basis for disruptive, new products, services and business models.

Our promise: as a strong solutions partner, we contribute our in-depth and broad industry expertise as well as our unique cross-industry innovation know-how, and take responsibility for the success of our clients' projects. We live for «empowering ideas».

Who we are

Zühlke is an independent service provider for product and software engineering, management consulting, and start-up financing. Zühlke provides added value as a result of the experience gained through more than 8000 successful inter-national projects, as well as continued investment in business and technology knowledge and understanding. Founded in 1968, the Zühlke Group today has local teams in Austria, Germany, Serbia, Switzerland and the United Kingdom. In 2016, Zühlke generated CHF 133 million in revenue, employing 800 staff.



Experience the Digital Economy

Our message is "Live Business Is Simple" and, with the new SAP S/4HANA suite, you can discover how you can connect your business processes in real time – across all departments.

SAP helps companies achieve the agility and flexibility they need to meet the challenges of the digital economy.

As well as demonstrating technologies to evolve IT infrastructure, we show you use cases in mission-critical topics such as the Internet of Things, customer experience, analysis and decision making, and collaboration with suppliers and service providers. We invite you to reinvent your company with solutions from SAP.

With SAP Predictive Analytics, business users, analysts, and data experts can use the findings from predictive analyses and forecasts. They can also automate the development and management of forecast models. The insights and forecast results gained on the basis of real-time data facilitate sound and lucrative decision making across the company.

As the market leader in enterprise application software, SAP (Schweiz) AG helps companies of all sizes and in all industries run better. From back office to boardroom, warehouse to storefront, desktop to mobile device – SAP empowers people and organizations to work together more efficiently and use business insight more effectively to stay ahead of the competition. SAP applications enable around 300,000 customers to operate profitably, adapt continuously, and grow sustainably. For more information, visit http://go.sap.com/product/analytics/predictive-analytics.html



Balzano has more than 25 years of experience in offering A.I. based commercial software solutions.

With its headquarters in Zurich, Switzerland, the company is actively engaging in industry/research partnerships with public and private institutes, such as ETH Zurich and

Microsoft Research.

We leverage Cloud technology to provide solutions more quickly and in a more agile manner than traditional IT. Our solutions integrate publicly available cognitive services, thus keeping costs low. They carry the official "Swiss Made Software" label.

Our offering includes custom-built projects as well as vertical solutions for healthcare, public safety, emergency aid and retail.

ScanDiags

Medical imaging plays an essential role in the diagnosis and treatment of diseases and health conditions.

The number of MRI, CT and other scans produced every year grows steeply, with the number of radiologists lagging increasingly. The image-based diagnostic process can only be performed by highly skilled experts. Still, it contains elements of repeatable work not related to the ultimate diagnosis, thus blocking time of these expensive specialists from the more important diagnostic process.

Furthermore, the knowledge- and experience-sharing process between radiologists and attending doctors typically uses time-consuming methods such as attending conferences and seminars, or writing and reading studies and reports.

Through artificial intelligence, built into its product ScanDiags, Balzano helps reducing the cost and increasing the benefit of the interpretation of medical images.

ScanDiags assists radiologists in analyzing orthopaedic MRI scans, reducing the time required for the typical diagnostic process. The product lists structures that can be detected in MRI sequences. It produces the raw structure of a report that can then be completed with further insights and a diagnosis by the radiologist or doctor.

ScanDiags also processes feedback that is submitted from finalized medical reports - and continues learning. An instance of ScanDiags represents the central repository of experience and methodology of any organization. It allows time savings, knowledge transfer, second opinion, retrospective quality-review, and further scenarios. Plus, ScanDiags never tires and works overtime at no additional cost.



SAS® Viya™: Quickly, easily solves prickly business challenges

Open and governed analytics help businesses build predictive models and agile machine learning applications. A cloud-ready, scalable, open platform for modern machine learning for all analytics skill levels, SAS® Viya™ extends and com-

plements SAS®9 with new and updated applications. Business analysts, data scientists and software developers can simultaneously derive insights from the same big, in-memory data creating portable analytical assets to solve their most challenging business problems. Through public application program interfaces (APIs) and support for multiple programming languages, SAS Viya extends capabilities in interactive discovery and reporting, statistics, data mining, machine learning, streaming data analytics, forecasting, optimization and econometrics.

SAS Viya is more than SAS with a pretty, cloud-deployed, HTML5 web analytics face for broader audiences. It is a rethink of the classic SAS platform, and with its emphasis on machine learning, positions SAS to deliver what Ovum terms 'smart analytics. Recognizing that users have different experiences and skill sets when it comes to analytics, SAS Viya products are designed to meet their needs and drive results:

Business users – SAS Visual Analytics, deployed at thousands of customer sites globally, now runs on SAS Viya and offers a popular self-service analytics and data visualization solution. Multiple users can visually discover relevant relationships in data, create and share interactive reports and dashboards, and use self-service analytics to quickly assess probable outcomes to make smarter, data-driven decisions.

Business analysts – SAS Visual Statistics features the same, consistent visual interface as SAS Visual Analytics. It enables interactive data exploration and building of descriptive or predictive models. By using SAS Visual Statistics, business analysts can more easily collaborate with statisticians to quickly refine models and make better-informed decisions.

Data scientists – SAS Visual Data Mining and Machine Learning was developed for data scientists who want to easily apply machine learning and data mining techniques to structured and unstructured data. With the same, friendly, consistent visual and programming interface, data scientists can expedite model building and code generation that is portable and scalable across different deployment environments.

Intelligence analysts – SAS Visual Investigator, a new version that addresses a wide variety of intelligence analysis and investigation management needs by identifying patterns, unknown relationships and objects of interest. Analysts can author scenarios to detect anomalies and suspicious events, manage intelligently prioritized alerts and conduct targeted investigations leading to better, more streamlined decision making and optimized business processes.

SCIGILITY SOLUTIONS FOR BIG DATA

Who we are

Scigility is the leading expert for Big Data Solutions in Switzerland. Comprised of an eclectic, multilingual and international team of Big Data experts, Scigility works closely

with industry as well as academic partners to remain on top of the newest developments in the field of Big Data.

Based in Zurich, active all over Switzerland as well as abroad, Scigility has more than tripled in size since the incorporation in 2013.

We create pragmatic, out-of-the-box solutions, troubleshoot and teach with passion, enthusiasm and precision.

What we do

We encourage our clients to take the next step in digital evolution by offering comprehensive services focusing on Big Data that are divided into four pillars:

- · Architecture & Conception
- Development, Analytics & Reporting
- Installation, Operation and Instruction
- Legal & Governance



How we do it

We enable our clients to understand, use and develop their digital strategy with Big Data technology within a national and international legal and governance framework.

We empower our clients from conception to governance – by teaching, supporting and developing state of the art platforms that fits into a global corporate IT, Business and Legal concept – we leave nothing out.

We build bridges in a digital world, with our experts we aim to optimize the transition to Big Data technologies by taking a big picture approach that goes beyond implementing new technology.





Innovative solutions spanning the real estate life cycle

Wincasa is the leading real estate service provider in Switzerland. With our team of 820 specialists, we offer our customers an extensive service portfolio covering the entire lifecyc-

le of properties, from initial planning, construction and management to revitalisation and sale of the property. The public limited company founded in 1999 with its head office in Winterthur is part of the Swiss Prime Site Group.

Wincasa is perceived as the leading competent and strategic advisory company, based on its broad spectrum of internal expert know-how, proactive development and expansion of knowledge as well as application of state-of-the-art technologies and security standards.

Playing a pioneering role in digitalisation

Digitalisation is fundamentally changing the real estate sector. Wincasa recognised at an early stage the necessity for digital transformation and corresponding change management. As sector leader, the company has been increasingly focusing on digital leaders from other sectors, systematically pursuing digitalization and developing the digital skills of its employees.

In order to maintain and extend its competitive capabilities, Wincasa is executing numerous projects within the scope of its digital strategy revolving around mobility, communications, process efficiency and data utilization. Well integrated real estate platforms offer various opportunities for all participants. The latest project comprises the «My Wincasa» communications portal, with which tenants are able to manage everything relating to rental agreements easily, electronically and without the need of any paper.

Community Sponsors

Swiss Statistical Society



The Swiss Statistical Society (SSS, www.stat.ch), founded in 1988, propagates application and development of statistics in Switzerland, represents the interest of professionals working in this field in relation to practice, research and education and contributes to the recognition of statistics as a scientific discipline in its own right. It fosters contacts between statisticians in administration, business and institutions of research and education. It supports cooperation between all institutions which deal with such goals.

The Society

- edits a bulletin three times a year. The bulletin informs about the activities and upcoming events in statistics in Switzerland.
- organizes the yearly Swiss Days of Statistics.
- consists of three Sections (official statistics, education / research and business / induestries).
- organises short courses oriented towards practical applications.
- supports a yearly seminar of PhD students.

Swiss Association for Analytics



The main objective of the Swiss Association for Analytics is to spread the word about analytics in Switzerland. Our goal is to show added value of predictive analytics, data mining and machine learning to Swiss companies.

To exchange ideas, share news and job offers, we manage a LinkedIn page (www. linkedin.com/groups/4586163). To facilitate networking, we organize regular Meetups (www.meetup.com/swiss-analytics). To share expert knowledge, we publish the Swiss Analytics Magazine (online and printed). You can find it, and many more content, on our website: www.swiss-analytics.com

SGAICOCo-organisers

Artificial Intelligence and Cognitive Science Special Interest Group of the Swiss Informatics Society



- Promotes intelligent technologies for innovation in our society
- Provides a platform for exchange between industry and universities
- Is open for your ideas and initiatives

SGAICO events

- · Bring together members that work on methods and technologies
- Discuss and disseminate related knowledge
- Explore interdisciplinary contexts such as for example engineering, medicine, psychology and law
- Establish contacts between users and experts in Switzerland to exchange on applications
- Promote education in Switzerland

Join the largest and fastest growing network for Al in Switzerland to discuss and experience the technology everybody talks about!

More Information: https://sgaico.swissinformatics.org/

Contact: sgaico@swissinformatics.org





Swiss Alliance for Data - Intensive Services Organisers



Who we are

The Swiss Alliance for Data-Intensive Services, Data+Service, is a nation wide network of competence from industry and academia focusing on the topic of data-intensive products and services. It intends to

strengthen the position of the Swiss economy by concrete activities:

- · Connecting and inspiring innovators and experts from industry and academia
- Enabling research and innovation projects in collaboration of industry and academia
- Developing expertise and knowledge in Switzerland by providing education and supporting recruiting

Motivation

The market of data-intensive products and services is especially fast growing: Joining and analyzing data intelligently creates new business opportunities in almost all industry sectors. To turn this opportunity into a successful business, two elements are necessary:

- 1. New insights have to be generated by connecting and analyzing data
- 2. These insights have to be converted into new marketable products and services (data products)

Activities

For promoting data-intensive services, Data+Service engages in various activities in four different areas:

- Education and professional development: Courses and educational programs for professional development in data science and service science
- Creating excellence in practice: Creation of "Expert groups": connecting and developing professionals in specific areas
- 3. Innovation and research: Launch of collaborative research and innovation projects between academia and companies to develop new data products and services
- 4. Creating a data-intensive community in Switzerland

Get in touch

Find more information on the web: http://www.data-service-alliance.ch/ Get in touch with us directly:

- Prof. Dr. Christoph Heitz, President, Data+Service Alliance, christoph.heitz@data-service-alliance.ch
- Prof. Gerold Baudinot, Managing Director, Data+Service Alliance, gerold.baudinot@data-service-alliance.ch
- Follow us on Twitter: @DataServiceCH

Conference Parallel Sessions

Room: Bellavista 2

Support for Research Projects & Start-ups

Financial support is one of the primary concerns/objectives while formulating a research proposal. We bring together at this platform, experts from Commission for Technology and Innovation (CTI) and EUResearch who will discuss the various funding opportunities for research projects, both through Swiss national funding as well as at the European level.

EUResearch will give information about how to prepare an EU R&D proposal, forming a consortium, what calls are currently open and more. CTI will talk about how their funding works, CTI's expectation with regard to the proposal content, innovation and the business aspect for the Swiss industry involved in the project. CTI will also give information about their support programs for start-ups and spin-offs. Join this session to hear from and meet the experts.

Time	Talk details
11:00	EUResearch - Funding opportunities in EU research projects
11:15	CTI projects - How they work
11:30	CTI Support offer for Start-ups & Spin-offs
12:00	Roundtable discussion - Startup

Use-Case Talks

An interactive session where companies come ahead with their use-case problems and seek answers from you. In the individual use-cases, will be discussed new technologies and applications where the presenter is facing technical problems. The presenter will explain what they have developed and where they stumbled into problems and you will provide them with answers with your expertise in that specific field. Please look at the description of the use-cases to see what ignites your interest and sign-up to participate.

Time	Use-Cases	
13:45	Helsana	eBay
14:30	SWITCH & ZHAW	Valora
15:15	PwC	

Helsana - 13:45 - 14:30

Room: Bellavista 2 Use-Case Talks

Building an Analytics Infrastructure – the right questions to ask

Building a hardware infrastructure for Analytics is a very challenging endeavor. Coming from a fragmented application landscape with Linux, windows, single- and multi-tier architectures makes the switch much more difficult and complex. Many models and batch-jobs requires up-to-date data and therefore there are strong dependencies between different worlds (for example data warehouse and Analytics) that need to be considered and planned for. The number of teams involved makes even more difficult to define and decide on a target architecture, since everyone have its own different requirements. Sizing (RAM, CPU, HD), number of environments and tools (and their integration) play a fundamental role for successful analytics infrastructure. There are fundamental questions that needs to be clarified before deciding on the target infrastructure: how will the deployment occur? How many environment are needed? What is exactly is a development environment since the data scientists will always need productive data? What kind of data sources you have (Oracle, Teradata, files, etc.)? How many? What is more important CPUs or RAM? What is the role of IT and of Data Scientist (does such a division make sense)?

In this use-case we will share our experience and what are the challenges and the right questions to ask before defining a target Analytics Infrastructure.

In addition we would like in particular to share our experience on the distinctions between the role and responsibilities of "IT" and "Data Scientists". Does such a distinction

Questions

make sense?

- What is an analytic Infrastructure? Are we talking only about hardware?
- How many environment are needed? Is a "Test" environment needed?
- What is the role of a "development" environment?
- What role the deployment process play in defining the target infrastructure?
- How important is the integration of different tools (for example R and SAS)? Is needed?
- How Analytics projects differentiate itself from "standard" IT projects (for example from - Waterfall and Agile methodologies)?
- What are the roles and responsibilities of IT and of Data Scientists? Does a distinction make sense in our data driven world?

Target audience

We recommend IT people, Data Scientist and Analytics Architects (and decision makers) to attend to gain insights of what it means to consolidate a complex infrastructure coming from a fragmented one and ensure success.

eBay - 13:45 - 14:30

Room: Bellavista 2 Use-Case Talks

Identification of Experts Possible Approaches

eBay is a unique marketplace with millions of live items and millions of active users. Naturally, with such a massive ecosystem there is considerable diversity in both quality of inventory and user expertise / knowledge. This is especially true in categories with specialized equipment (e.g. cameras, snowboards, mountain bikes).

The ability to identify users with expertise in a domain or high-quality inventory would enable us to provide interesting recommendations, give us deeper inventory insights, etc.

We have many data points related to user browsing and purchasing behaviour that we can leverage to identify expertise and quality. We have implemented a modified version of the graph-based SPEAR algorithm (Noll and Yeung 2009) for this purpose. However, we have open questions and are open to other approaches.

Questions

How would you build a model to identify

- · high-quality inventory and
- domain experts at scale?

Target Audience

Data science practitioners / students interested in applied machine learning. Those with experience in graphical models are encouraged to attend.

SWITCH & ZHAW - 14:30 - 15:15

Room: Bellavista 2 Use-Case Talks

Automated provisioning of data analytics clusters

Data analytics framework such as Hadoop, Spark offer very powerful tools to analyzing huge amount of data and gain deep insights. But managing a reasonably optimized cluster is complex. With an aim to remove the complexity and automate the lifecycle management of these clusters for research groups and the student community within Switzerland, ZHAW and SWITCH have been collaborating within the SCALE-UP project. The researchers at ICCLab in

InIT, ZHAW have developed DISCO - a distributed computing framework orchestration solution which is able to offer creation of on-demand clusters of Hadoop, Spark in a few minutes on SWITCHengines. Additional tools such as Zeppelin and Jupyter can also be provisioned with just a click through DISCO. But there still remain a few questions to be answered, which will be addressed in the discussion.

Questions

- How big are typical big-data use cases in the various industries in terms of storage needs?
- SWITCHengines is a cloud solution primarily for the Swiss academic and research sector, developed and managed by SWITCH. The data is hosted in Switzerland where the Swiss data protection law applies. What are further essential requirements which will make the data analytics solution offered via DISCO over SWIT-CHengines desirable (apart from cost factors) when compared to commercial providers located outside of Switzerland / Europe?
- Are organizations use cases mainly representative of long lived analysis tasks, or are organizations and groups open to using a cluster only for the duration of the tasks and then deleting or pausing the cluster resources to minimize costs in the long run?
- The clusters created by DISCO come with their own dedicated storage which is not shared with any other clusters as a default. Is this assumption a reasonable one, or is there a general consensus in the user community that data should be more easily shared between different clusters?
- Apart from Hadoop and Spark, which are the other non-commercial popular frameworks used by the community?

Target audience

The presentation itself is meaningful for anyone working with open source data analytics frameworks such as Hadoop, Spark - students of big data and allied subjects such as AI and machine learning are suitable audience. Companies offering Business Intelligence solutions, big data cluster administrators from universities and companies are encouraged to attend.

Valora - 14:30 - 15:15

Room: Bellavista 2 Use-Case Talks

Demand forecasting in high-frequency retailing

Data Science at Valora is part of Valora's LAB focused on building a superior shopping experience by developing new digital products and innovative services. Understanding our data is key in this process. The LAB's analytics team accelerates business by providing analytical insights to everyone in the company.

A particular cornerstone in the multiverse of analytics tasks is demand forecasting to streamline business processes at POS level.

Our use case concerns the Valora kkiosk brand with more than a thousand POS in Switzerland. The goal is to develop better and new prediction models that further improve the efficiency in our supply chains and order systems. We look forward to discuss aspects of this journey with you in the session.

Questions

- Choosing the right model/s for demand forecasting, univariate vs. multivariate.
- Analysis of computational and maintenance costs.
- Usability and scalability of a fully-fledged high-dimensional multivariate model.
- Making POS-tailored forecasts including POS specifics, assortment cubes, external factors, and so on.
- Sourcing external data in a consistent and persistent way, e.g., weather data; also w.r.t. granularity, and data quality.

Target audience

Data scientists and researchers (students and professionals) with interest or experience in the field of demand forecasting and model building.

PwC - 15:15 - 16:00

Room: Bellavista 2 Use-Case Talks

Machine learning in transactional data

At PwC, data are at the heart of everything we do. We have a wealth of experience in all data-related disciplines from collection, cleansing and management to building analytical algorithms and visualisation tools. We combine this expertise with other PwC's professional offerings such as Tax Services, to create powerful, data-enabled solutions for clients.

We present a use case resulting from the application of Machine Learning to tax statements of corporations. Taking this case as a basis, we elaborate on other use cases in various business settings. In a group discussion during the use case session, we encourage you to share own challenges in your organizations with the group and to take home practical insights for your organization.

Questions

- What are ways to build machine learning value chains in environments with high requirements in accuracy and precision?
- What experience do you have with artificial intelligence in any kind of transaction related problems?
- Have you looked into technology for coping with regulations in your own organizations?
- Have you checked possibilities of automating office processes (aka Robotic Process Automation) and may you share your experiences?
- If you haven't had the chance to deal with those questions, what the obstacles in your environment?

Target audience

Students and professionals with interest in accurate, transaction-style data, such as accounting, tax and regulatory.

Notes

Program Committee

Prof. Gerold Baudinot, Data+Service Alliance, ZHAW

Prof. Dr. Martin Braschler, InIT, ZHAW

Dr. Mark Cieliebak, InIT, ZHAW

Dr. Oliver Dürr, IDP, ZHAW

Dr. Jörg Gerigk, PwC

Stefan Hegyi, SML, ZHAW

Prof. Dr. Christoph Heitz, IDP, ZHAW

Melanie Imhof, InIT, ZHAW

Dr. Michael Luggen, eXascale Infolab

Prof. Dr. Andreas Ruckstuhl. IDP. ZHAW

Prof. René Schumann, HES-SO

Dr. Christian Spindler, PwC

Dr. Thilo Stadelmann, InIT, ZHAW

Prof. Dr. Kurt Stockinger, InIT, ZHAW

Dr. Michael Zehnder, GateB

Dr. Marcel Zemp, D ONE

Organization Committee

Gerold Baudinot, Melanie Imhof, Amrita Prasad

SDS|2018

We hope you enjoyed SDS|2017 and got useful insights and contacts. Please use the feedback forms to help us optimize SDS|2018. You can also send your comments and ideas for participation to

gerold.baudinot@data-service-alliance.ch

Thank you!

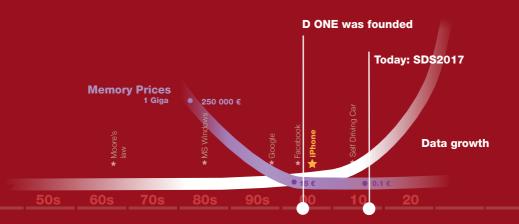
Imprint

Swiss Alliance for Data-Intensive Services c/o Prof. Gerold Baudinot ZHAW, Gebäude TD Postfach, 8401 Winterthur Phone: +41 58 934 70 68

Email: gerold.baudinot@data-service-alliance.ch

D ONE: READING THE GAME. SINCE 2006.

Why digitization is important today.





Data Driven Value Creation

