SDS2019 + SWISS CONFERENCE ON DATA SCIENCE

June 14, 2019

SDS2019 – Inspire. Interact. Innovate.

The Swiss Conference on Data Science is the day and place where science, innovation and business unite. The conference offers an in-depth exploration of new insights, techniques and methods, relevant for Switzerland and its market players.

Get inspired by excellent keynotes, participate in interactive sessions, and network with professionals to plan your next Data Science based innovations.

data service

Swiss Alliance for
Data-Intensive Services

FOREWORD

Welcome to SDS2019

This is the 6th SDS conference organized by the Swiss Alliance for Data-Intensive Services (Data+Service), and it keeps growing. Based on high demand, the conference is now organized in a scientific and a business stream.

The conference program is packed (a big thank-you to the program committee): 50 slots tackling key topics from AI management to data ethics - an excellent place to learn from each other and to exchange ideas.

We wish you a rewarding conference, interesting encounters, exciting insights and great inspirations at SDS2019. 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

Swiss Alliance for
Data-Intensive Services

		TRACK 1 @ARENA	TRACK 2 @SOPRA GRANDE	TRACK 3 @SOPRA 2-3
6	9:30	Welcome		
	9:45	Keynote 1: Ken Hughes on Data²: Where Silicone Meets Soul		
	10:30	Break		
6	11:00	Deep Learning in Practice	Natural Language Processing	Transportation Applications
8	11:30	Deep Learning in Practice	Natural Language Processing	Transportation Applications
10	12:00	Deep Learning in Practice	Natural Language Processing	Transportation Applications
12	12:30	Deep Learning in Practice	Natural Language Processing	Transportation Applications
	13:00	Lunch & poster session		
14	14:00	Machine Learning Applications	Natural Language Processing	Machine Learning Methods
16	14:30	Machine Learning Applications	Natural Language Processing	Machine Learning Methods
	15:00	Break		
18	15:30	Machine Learning Applications	Natural Language Processing	Machine Learning Methods
20	16:00	Machine Learning Applications	Natural Language Processing	Machine Learning Methods
	16:30	Keynote 2: Aleksandra Przegalinska on Context-Aware AI Systems of t		
20	17:15	Award Ceremony & Closing		
	17:30	Apero riche & poster session		

TRACK 4 @BELLAVISTA 5	TRACK 5 @BELLAVISTA 2-3	TRACK 6 @ORIONE	TRACK 7 @BELLAVISTA 4
Data Science Operationalization	Spatial Data Analytics	Sponsored Talks	
Data Science Operationalization	Spatial Data Analytics	Sponsored Talks	
Data Science Operationalization	Spatial Data Analytics	Sponsored Talks	
Data Science Operationalization	Spatial Data Analytics	Sponsored Talks	
Data Science Operationalization	Data Ethics	Sponsored Talks	
Data Science Operationalization	Data Ethics	Sponsored Talks	Sponsored Talks
Roundtables	Data Visualization	Sponsored Talks	Data Ethics
	Data Visualization	Sponsored Talks	Data Ethics

he Future

Welcome & Keynote Speaker - 9:30am - 10:30am

Welcome

Aileen Zumstein - Aizu Communication

Keynote - Ken Hughes

Data²: Where Silicone Meets Soul

Artificial Intelligence is the next significant consumer step-change. In the late-90s and early 2000s it was the Internet and eCommerce, from 2007+ it was the smart phone and mCommerce. 2020+ will be the decade where AI changes the nature of business, consumerism and society once again. While B2C and B2B are established norms, B2M (Business to Machine) will fundamentally change the nature of how we buy and live our lives. A future run by data and algorithms, with low-to-medium involvement purchases undertaken by our virtual assistants, a future where the 'consumer' may not be as involved as before in the consumer decision making process. This goes beyond the usual data ethics 'personalization V privacy' argument. This is about cultural and societal change. A world where the silicone runs our lives more than ever before. But as always, that which will become scarcer will become more valued, and as such, genuine human interaction and customer experience will become even more of an asset. The businesses that build for this 'Silicone Meets Soul' future will be the ones that succeed, balancing the instant, seamless, personalized and frictionless delivery via the 'machine' with the human feels of a business that cares. Always remember that data is useless unless activating it makes us feel something as a consumer.

Breakout Session 1 - 11:00am - 11:30am

Track 1: Deep Learning in Practice
@ARENA

Deep Learning Solutions for Cost-effective Tunnel Maintenance

Every year 4700km of new tunnels are built, amounting to a total annual growth value of 7%. Until today, tunnel assessment is mostly based on a slow and subjective human inspection process. As the total length of tunnels which must be inspected is constantly increasing, it becomes more and more important to guarantee their operation reliability with safe and cost-efficient means.
Amberg Technologies has developed in partnership with LeanBI a new platform for tunnel inspections based on Artificial Intelligence. The new platform achieves higher degree of automatization, by applying deep learning models for image segmentation.
The talk will explain the principles of the new platform, and then focus on the implications and results of deep learning modeling applied for automatic tunnel evaluation.

Vassilis Kalofolias, LeanBI
Edouard Lamboray, Amberg Technologies

Track 2: Natural Language Processing
@SOPRA GRANDE

Resilient Combination of Complementary CNN and RNN Features for Test Classification through Attention and Ensembling

State-of-the-art methods for text classification include several distinct steps of pre-processing, feature extraction and post-processing. In this work, we focus on end-to-end neural architectures and show that the best performance in text classification is obtained by combining information from different neural modules. Concretely, we combine convolution, recurrent and attention modules with ensemble methods and show that they are complementary. We introduce ECGA, an end-to-end go-to architecture for novel text classification tasks. We prove that it is efficient and robust, as it attains or surpasses the state-of-the-art on varied datasets, including both low and high data regimes.

Athanasios Giannakopoulos, Swisscom

Track 3: Transportation Applications
@ SOPRA 2-3

PHom-GeM: Persistent Homology for Generative Models

Generative neural network models, including Generative Adversarial Network (GAN) and Auto-Encoders (AE), are among the most popular neural network models to generate adversarial data. However, generative models are known to provoke chaotically scattered reconstructed distribution during their training and incomplete generated adversarial distributions. Current distance measures fail to address this problem. We propose Persistent Homology for Generative Models, PHom-GeM, to assess and measure the distribution of a generative model. PHom-GeM uses persistent homology, the study of the topological features of a space at different spatial resolutions, to compare the nature of the true and the generated distributions. Our experiments underline the potential of persistent homology for Wasserstein GAN in comparison to Wasserstein AE and Variational AE.

Jeremy Charlier, University of Luxembourg

Track 4: Data Science Operationalization
@BELLAVISTA 5

Democratizing Data @ Mobiliar to Foster Innovation

One of the biggest challenges that data science currently faces is the limited availability and access to relevant data. This problem is particularly acute within large enterprises, where the expertise of data scientists cannot be fully exploited because rigid IT structures hinder their ability to locate the appropriate data assets for their use cases. To successfully accelerate the digital maturity and the potential of innovation in enterprises new ways of data organization have to be found. SPOUD created the Data Market, an enterprise data integration platform based on Apache Kafka that allows data scientists, data engineers, domain specialists and business people alike to access, understand and use all available data assets in real time as data streams or full states. In this session we are going to showcase how the Swiss Mobiliar is building a data driven culture supported by tools like the Data Market. Using a concrete use-case inside the Mobiliar we are going to explain how people and technology have transformed to create faster and more efficient data science feedback loops.

Matthias Rüedlinger, SPOUD
Bruno Russiniello, Mobiliar

Track 5: Spatial Data Analytics
@BELLAVISTA 2-3

LoRaLoc: Machine Learning-Based Fingerprinting for Outdoor Geolocation Using LoRa

LoRa technology allows long-range transmissions with low power consumption and it can also be used indoor. For these reasons, the introduction of a precise time-stamping of LoRa frames provides the possibility to use this technology for accurate localization in many scenarios. However, this is still very challenging to achieve in non-line-of-sight environments such as urban landscapes. In this paper, we present a “fingerprinting” method to perform outdoor geo-location based on machine learning (Random Forest and Neural Networks) applied to a reference map. The map combines Time Difference Of Arrival (TDOA) measurements generated by a LoRa network and GPS location as ground truth. We tested our approach on simulated data achieving promising results with a Root Mean Squared Error below 9 meters by using a Long Short-Term Memory (LSTM) network.

Francesco Carrino, University of Applied Sciences and Arts Western Switzerland

Track 6: Sponsored Talks
@ORIONE

Real-Time Learning and Prediction in (Un) structured Data

We present a generalized machine learning framework for learning spatially and temporally in sequential data. While most machine learning techniques rely on dense connections of information (e.g. neural networks) that rely on derivatives, back-propagation, batch training, and regularization, rendering them incapable of learning in real-time, we show that machine learning can be done in real-time using sparse distributed representations (SDR). The talk will discuss the history of SDR and propose a formulation for learning on all kinds of sequential data. We conclude with some Use Cases that are currently being used in various industries, from music, to anomaly detection in financial data.

Christian Blakely, PwC

Breakout Session 2 - 11:30am - 12:00pm

Track 1: Deep Learning in Practice
@ARENA

Deep Learning for Retina Segmentation: from Prototype to Production

Convolutional neural networks have proven their capability for image segmentation in medical settings. In this talk, we present the development and deployment of a cloud-based, deep learning algorithm for the semantic segmentation of Optical Coherence Tomography (OCT) images. OCT is an imaging technique yielding 3D-images of the human eye. Our convolutional neural network that segments retinal tissues was trained with selected ground truth data and evaluated against benchmark segmentations. During the production step, we made the trained network accessible through a web interface and deployed the data processing back-end to a cloud server. The presented results have been achieved in collaboration between Supercomputing Systems (SCS) and Dr. med. Peter Maloca, Institute of Molecular and Clinical Ophthalmology Basel (IOB).

Pascal Kaiser, Supercomputing Systems,
Susanne Suter, Supercomputing Systems

Track 2: Natural Language Processing
@SOPRA GRANDE

APCNN: Tackling Class Imbalance in Relation Extraction through Aggregated Piecewise Convolutional Neural Networks

One of the major difficulties in applying distant supervision to relation extraction is class imbalance, as the distribution of relations appearing in text is heavily skewed. This is particularly damaging for the multi-instance variant of relation extraction. We introduce a new model called Aggregated Piecewise Convolutional Neural Networks, or APCNN, to address this problem. APCNN relies on the combination of two neural networks, a novel objective function as well as oversampling techniques to tackle class imbalance. We empirically compare APCNN to state-of-the-art approaches and show that it outperforms previous multi-instance approaches on two standard datasets.

Alisa Smirnova, University of Fribourg

Track 3: Transportation Applications
@SOPRA 2-3

How Expensive is Punctuality? Using Data to Manage Track Maintenance Projects for SBB

The Swiss federal railways (SBB) runs one of the most complex infrastructures in Europe, under very high expectations from the public and the government. But even Swiss train tracks have a finite life span, and unsurprisingly, doing maintenance on a track network of over ~3200km while keeping things running smoothly is a costly undertaking.

SBB started digitizing many aspects of these maintenance projects years ago and ultimately the vision is to use data and advanced analytics throughout the whole process -- from project selection and specification, to project scheduling and review -- and giving all the involved stakeholders easy access to the wealth of knowledge already accumulated, and how this can improve efficiency and save costs.

The journey towards this vision has been to iterate over small well-defined problems and leverage the results for the next steps, as well as to integrate solutions quickly into production. In particular, we have developed an automated tool for generating suggestions to reduce costs on project-level, whereas on portfolio-level, we have created a way to quantifiably assign costs to interrelated cost drivers.

Axel Obermeier, D ONE
Daniel Auvinet, SBB

Track 4: Data Science Operationalization
@BELLAVISTA 5

Principles and Best Practices for Applied Machine Learning Models (in Industry)

Data science is maturing as a field, since interest increased rapidly in recent years. Many organizations are beginning to consolidate their activities around data analytics, model building and productive deployment, and are in the process of professionalizing these activities. Still other organizations are taking the first steps on their journeys into data science, and would like to learn from those who have gone before.

At Swiss Re, we brought together the collective expertise and experience of numerous

expert practitioners and managers from data science and risk management to create a definitive set of principles and best practices that guides all our data science activities.
In this talk we present and discuss these principles and emphasize the principles which need much more care by the Data Scientist in industrial applications than in education and research.
Mark Rowan, Swiss Re

Track 5: Spatial Data Analytics
@BELLAVISTA 2-3
Analyzing the Geo-Spatial Coverage of Veterinary Services
Dogs and cats belong to the most popular companion animals in Switzerland. Accordingly, a dense network of veterinary practitioners guarantees these pets' well-being and health. From an economical perspective the total distance between all pets and their respective vet should be minimized.
This objective was the trigger for a joint project between the Alliance members identitas and geo7. The aims of the project are to collect and pre-process the required data basis, derive spatial statistics and patterns, visualize regional differences in the supply situation and draw conclusions for the relevant local, regional and national actors.
The core of the analysis is a specific gravity model based on drive time distances. The project goals were to choose the most appropriate formulation, adjust it based on empirical data and apply it nation-wide.
Nicolas Lenz, geo7

Track 6: Sponsored Talks
@ORIONE
Classification at PostFinance's Input Management
PostFinance, one of Switzerland's leading retail financial institutions, distinguishes around 360 business processes/cases (e.g. opening/closing accounts or address changes) that customers initiate by manually filling in forms and sending them physically back to PostFinance. So far, PostFinance employees subsequently scanned the received customer documents and assigned (classify) a part of them manually to a business process.
ELCA has developed a solution for automatic text classification that allows PostFinance to process physical customer documents more effectively. The customized and flexible solution with a classification accuracy of 80% significantly increased the processing rate. The solution is fully integrated into the existing IT landscape at PostFinance.
Robert Gwadera, ELCA Informatik
Dennis Furrer, PostFinance

Breakout Session 3 - 12:00pm - 12:30pm

Track 1: Deep Learning in Practice
@ARENA
SmartProfile – Deep Learning for the Powder Coating Industry
Goal of powder coating is to build an anti-corrosive and esthetical 'skin' on the profile. One of the challenges is the unique identification of profiles. There exist more than 4'000 profiles which vary sometimes only by a few millimeters. How to make sure that for each order the right profiles are delivered to the machine? Industrial solutions like barcode, QR-code, RFID tags do not work on a surface which will first be sprayed with powder and then cured at 200°C.
The envisaged solution is to use a smart phone, take a picture, process the image and provide the output directly on the smart phone. Key is a very robust deep neural network which has to deal with different smart phones manufactures and their cameras, variable distances, many backgrounds and all kind of illumination conditions.
Philipp Schmid, CSEM

Track 2: Natural Language Processing
@SOPRA GRANDE
Building a Data Culture with Veezoo
Peter Drucker said that "if you can't measure it, you can't improve it." - but how do you improve that what you've measured?

That's the key question and, in order to answer it, you need to go beyond just KPIs. You need to use data strategically, to come up with the best decisions. For that, data must be easily accessible to every employee. How can you make this possible?
Veezoo is the world's most intelligent virtual assistant to help your sales team crush their targets. It empowers them with the insights that they need to keep their team on track, increase sales and succeed with their customers. Veezoo is being used by one of the largest insurance companies in the world, AXA, building the data-driven culture that 21st century companies need to be market leaders.
Marcos Monteiro, Veezoo

Track 3: Transportation Applications
@SOPRA 2-3
Tapping into New Risk Pools: Leveraging Machine Learning to Create a Flight Delay Parametric Cover
Flight delay insurances already exist, but are often bundled into broader travel insurances that are expensive, have cumbersome claiming processes and in-transparent exclusion rules. Looking at this inefficiency, Swiss Re created a white-label flight delays insurance product offering a lean, easy-to-buy and fully automated product from the underwriting to the claim settlement.
Naturally, frequent travelers have strong views on which flights might be delayed, and anti-selection is therefore high. To mitigate it, the solution relies on a dynamic pricing engine which is estimating the propensity to delays of any individual flight in the future, up to one year ahead. This pricing engine is fully leveraging machine learning techniques, joining historical and future flights and weather data.
Grégoire Caro, Swiss Re

Track 4: Data Science Operationalization
@BELLAVISTA 5
Improving Manufacturing Plants through Big Data Analytics
Data is pervasive and covers every aspect of our everyday life. But how do you turn piles of data into actionable insights that positively impact your manufacturing plant – and ultimately your business? In this talk we will present our project for a Swiss food manufacturer who decided to embark on a journey where flour and sugar meets data & machine learning.
We will start off by highlighting the specific problem that was tackled in the project: optimizing efficiency of the manufacturer's most important production line which runs 24/7 but consistently falls short of production expectations due to a high number of very short production interruptions.
We will then introduce our big data architecture blueprint that is adaptable to the needs of specific manufacturing requirements and is capable of predicting short production interruptions within five seconds of data ingestion. Through the use of machine learning,

we were able to find about 90% of these short interruptions and to identify the most relevant factors of influence that lead to a higher frequency of production interruptions.
To wrap up, we will share our lessons learned, i.e. with regards to where some of the emerging big data technologies fall short of expectations and how our proposed big data architecture can serve as an example for other companies that plan to build their solutions in a cost-efficient manner using cloud technologies.
Marc Schöni, Microsoft
Martin Weber, Noser Engineering
Kurt Stockinger, Zurich University of Applied Science

Track 5: Spatial Data Analytics
@BELLAVISTA 2-3
Revenue Forecasting and Store Location Planning at Migros
Every year, the ten regional Migros cooperatives build or rebuild many branches of their supermarket network. Due to the high financial impact of successful (or failed) alterations of the store infrastructure, accurate planning is crucial. To this end, we developed an algorithm that provides realistic revenue forecasts even for complex scenarios such as multiple new stores and/or alterations of the competitor store network. This algorithm combines a heuristic simulation of consumer behavior with machine learning methods to deliver both accurate and interpretable results. As the software has been developed in-house and in close collaboration with key users at Migros' ten regional cooperatives, we will also provide insights into the challenge of realizing complex data science projects while still providing sufficient interpretability for the planning expert.
Bojan Škerlak, Migros

Track 6: Sponsored Talks
@ORIONE
Machine Learning for Data Scientists Involving SAP Data Using Open Source Technologies
The aim of many Data Science projects is the improvement of a business process through Machine Learning. Such a business process often involves an SAP system. Data held in a SAP system might be required to train a model and derive to derive predictions. At the same time, the SAP system might be hosting the business process itself that benefits from the inference. However, the agile and empirical nature of a Data Science project can be in conflict with the operational needs of a mission critical ERP system.
This talk will explain the different options for Data Scientists to work with their familiar open source technologies such as Python, R or TensorFlow in an SAP infrastructure. SAP is also currently developing a new platform to support the handover of Machine Learning artifacts created by Data Scientists to the operational department for ongoing use. The talk will conclude by giving an update and outlook on this new platform that is aimed to bridge the gap between model creation and deployment.
Andreas Forster, SAP

Breakout Session 4 - 12:30pm - 01:00pm

Track 1: Deep Learning in Practice
@ARENA

Survey of Artificial Intelligence for Card Games and Its Application to the Swiss Game Jass

In the last decades we have witnessed the success of applications of AI to playing games. In this work we address the challenging field of games with hidden information and card games in particular. Jass is a very popular card game in Switzerland and is closely connected with Swiss culture. To the best of our knowledge, performances of AI agents in the game of Jass do not outperform top players yet. Our contribution to the community is two-fold. First, we provide an overview of the current state-of-the-art of AI methods for card games in general. Second, we discuss their application to the use-case of the Swiss card game Jass. This paper aims to be an entry point for both seasoned researchers and new practitioners who want to join in the Jass challenge.

Michele Alberti, University of Fribourg

Track 2: Natural Language Processing
@SOPRA GRANDE

Do You Have to Read All Incoming Documents?

A tax advisor has to read all receipts and fiscal documents of his customers and manually extract relevant information for the tax reports; a hospital easily receives hundreds of patients documents every day; and every customer support is overwhelmed with standard requests. In all these cases, the majority of documents are unstructured and are at present processed manually. This results in a tremendous amount of work, which can be – at least partially – automated using state-of-the-art tools from text analytics and machine learning.

One such solution uses automatic classification of incoming documents into pre-defined classes (such as medical reports, prescriptions etc.) and then extracting the relevant information for further processing. Using real-world examples, we provide an overview of the potential applications, a realistic assessment of the effort and the resulting quality to be expected.

Mark Cieliebak, Zurich University of Applied Sciences

Track 3: Transportation Applications
@SOPRA 2-3

FAIRTIQ: From Research Prototype to a Data-Driven Product

FAIRTIQ app allows people to obtain a valid public transport ticket with a single swipe in, and the best fare is computed automatically at the end of the trip. What makes the technology unique is that it was developed from a research study. This talk elaborates on challenges for transforming a prototype into a successful product, with a particular focus on making a solution robust for a variety of input data, whether a constant stream of new data is needed to make improvements, and what metrics need to be implemented to ensure high quality of the solution.

Roman Prokofyev, FAIRTIQ

Track 4: Data Science Operationalization
@BELLAVISTA 5

Data Science Infrastructure

The journey towards data-driven decision making usually starts out by integrating data silos into a DWH to facilitate reporting and enable building dashboards. Then the first 1-2 data scientists appear in the organization and start exploring this data, generate insights, and produce first ML models. These are usually run on the laptop or desktop of the data scientist for a couple of months. With their success, the team and the number of "productive" applications thrive - and the need for a data science infrastructure emerges.

We face such situations often with clients across industries and borders. Questions that arise are: How to put the model in production? How to deal with the trade-off agility of data scientists in the "lab" vs. stability of code in the "factory"? Can Python-/R-Packages be downloaded behind our firewall? Interfaces to remote machines? And what about management of versions, credentials, and releases, monitoring of model KPIs over time, and requirements for documentation, reproducibility, and auditability? Most importantly, how should the process organization be structured?

We will show how we addressed these questions for a team of data scientists at Raiffeisen.

Lukas Seger, Raiffeisen
Philipp Thomann, D ONE

Track 5: Spatial Data Analytics
@BELLAVISTA 2-3

From Terapixels to Crop Yields

The first proposals of exploiting remote sensing techniques to support the insurance industry date back to mid-eighties. Since then, thanks to the big growth of the remote sensing market, satellite data found their way into an increasing number of applications. Nowadays, with about 500 Earth observation satellites orbiting around our planet, the risk management support is more promising than ever before. Applying machine-learning techniques to the satellite information enables data-driven decisions. We are moving from pixel to the field, identifying cropped areas, which then are being classified and monitored, creating a dynamic crop mask. During that process, plenty of insights are gained out of the historical changes over time and space: risk factors, over performing farmers, market penetration tools, are some of the side products of such analysis.

Charilaos Tsarouchas, SwissRe

Track 6: Sponsored Talks
@ORIONE

Data Visualization ≠ Visual Data Analysis

Data is everywhere, but how do we make sense of it? Put it in front of us in big tables? What does that tell us about the data? Is seeing the data the same as understanding the data? This presentation talks about how visualizing data can help us identify and make sense of the information that's encoded in the data. It even goes one step further, when just visualizing the data and seeing it is not enough, when we want to go deeper and understand the data, find the insights, and derive meaningful actions from it. We will look at data visualization and visual data analysis in the past, what's possible now, and where it will move in the future.

Konstantin Greger, Tableau

Track 1: Machine Learning Applications
@ARENA

Automated Data Quality Assurance with Machine Learning and Autoencoders

Companies store massive amounts of data to derive business value from it. However, data quality issues limit the data's usefulness. Typical issues include missing values, wrong formats, or incorrect values. Usually companies try to implement hard-coded rules on the database entries to ensure their correctness. However, this approach is not universally applicable and does not scale.

Machine learning provides ways to significantly improve and speed up error detection, without the need to explicitly specify hard-coded rules. The secret is autoencoders – neural networks that model and reconstruct their own input. Adequately trained, they only reconstruct clean data, which exposes corrupted entries.

We applied autoencoders for data quality assurance on real-world data. Alongside the specific algorithms for different data types and our findings, we will present an R Shiny demo.

Milica Petrovic, InCube Group
Martin Müller-Lennert, InCube Group

Track 2: Natural Language Processing
@SOPRA GRANDE

Towards Reproducible Research of Event Detection Techniques for Twitter

A major challenge in many research areas is reproducibility of implementations, experiments, or evaluations. New data sources and research directions complicate the reproducibility even more. For example, Twitter continues to gain popularity as a source of up-to-date information. As a result, numerous event detection techniques have been proposed to cope with the increasing rate and volume. Although some works provide their implementation or conduct an evaluation, it is almost impossible to reproduce their experiments. The main drawback is that Twitter prohibits the release of any datasets that are used by researchers. We present a survey of the vast landscape of implementations, experiments, and evaluations. Furthermore, we propose a reproducibility toolkit including Twistor, which can be used to simulate an artificial Twitter data stream as input for the event detection techniques.

Harry Schilling, University of Konstanz
Andreas Weiler, Zurich University of Applied Science

Track 3: Machine Learning Methods
@SOPRA 2-3

Automated Machine Learning in Practice: State of the Art and Recent Results

A main driver behind the digitization of industry and society is the belief that data-driven model building and decision making can contribute to higher degrees of automation and more informed decisions. Building such models from data often involves the application of some form of machine learning. Thus, there is an ever growing demand in work force with the necessary skill set to do so. This demand has given rise to a new research topic concerned with fitting machine learning models fully automatically—AutoML. This talk gives an overview of the state of the art in AutoML with a focus on practical applicability in a business context, and provides recent benchmark results on the most important AutoML algorithms.

Lukas Tuggener, Zurich University of Applied Science

Track 4: Data Science Operationalization
@BELLAVISTA 5

Improving Reproducible Deep Learning Workflows with DeepDIVA

The field of deep learning is experiencing a trend towards producing reproducible research.

Nevertheless, it is still often a frustrating experience to reproduce scientific results. This is especially true in the machine learning community, where it is considered acceptable to have black boxes in your experiments. We present DeepDIVA, a framework designed to facilitate easy experimentation and their reproduction. This framework allows researchers to share their experiments with others, while providing functionality that allows for easy experimentation, such as: boilerplate code, experiment management, hyper-parameter optimization, verification of data integrity and visualization of data and results. Additionally, the code of DeepDIVA is

14

well-documented and supported by several tutorials that allow a new user to quickly familiarize themselves with the framework.
Michele Alberti, University of Fribourg

Track 5: Data Ethics
@BELLAVISTA 2-3
Towards an Ethical Code for Data-Based Business
In this paper, we outline the structure and content of a code of ethics for companies engaged in data-based business, i.e. companies whose value propositions strongly depends on using data. The code provides an ethical reference for all people in the organization who are responsible for activities around data. It is primarily targeted on private industry, but may also be used by public organizations and administrations. The code has been developed in a joint industry-academic initiative, involving specialists for ethics as well as for all relevant data-related issues.
Michele Loi, University of Zurich

Track 6: Sponsored Talks
@ORIONE
Digital Services in the Mobility Market: How to Optimize Value with Co-Creation
More people. More goods. Fewer resources. There's no end to the number of challenges facing rail operators today. The availability of fast, reliable and data-scientific information for railway operators will offer a sustainable alternative to aircraft and automobiles. With Digital Services no investment in expensive equipment is needed, with already available data from the assets we at Siemens develop the right services, mindset and culture in co-creation with our customers. Railigent from Siemens Mobility Services empowers customers to understand their rail data, generate valuable information and get more out of their systems.
Railigent makes intelligent use of rail data to create added value. Improvement of up to 100% availability – optimized maintenance and operation planning. Rail transport is made more efficient, reliable and safe thanks to condition-based monitoring, data analysis and predictive maintenance concepts. Co-creation using Siemens' domain knowledge and our clients' expertise allows for comprehensive asset management and customer proximity.
On the one hand, the presentation is aimed at all entrepreneurs who want to create value together with their customers and data science solutions and are looking for an approved business model. It is also aimed at data scientists and employees of railway operators who would like to get to know the Siemens solution and the Siemens way to handle Digital Services within our company. As A Head of Sales within Siemens, my task is to make intelligent use of rail data to create added value for my customers, the train operators in Switzerland.
Ruben André Lorenzo, Siemens

Breakout Session 6 - 02:30pm - 03:00pm

Track 1: Machine Learning Applications
@ARENA

Revenue and Availability Optimization Approaches for Shared Mobility Providers

A recognized challenge in emerging shared mobility services is the parallel optimization of availability and revenue. By setting the price dynamically for each transaction, a provider can react fast and directly to a changing environment, hence steer availability and revenue. Such a pricing strategy can be optimized stepwise by simulation price elasticity and demand with the help of data sources such as historic transactions, CRM, weather, public events and data from competing mobility providers. Ideal price variations which fulfill the maximization target are calculated through a reinforced learning algorithm, which is found to outperform explicit parameterized pricing agents. The newly developed AI-pricing approach enables, as first of its kind, to influence customer behavior in a predictive way for example by providing a low price to customers who bring the car to a location where the demand is expected to be higher in the near future. Hence, the algorithm optimizes revenue by forecasting the demand and availability of cars correctly in different time horizons and areas.

Daniel Müller, ITM Predictive
Jan Fränkle, ITM Predictive

Track 2: Natural Language Processing
@SOPRA GRANDE

Triangulated Sentiment Analysis of Tweets for Social CRM

Social media platforms like Twitter present an unprecedented opportunity for social customer relationship management by analyzing the ongoing discussions about business events such as a service outage. These opinions have been analyzed for sentiment with lexicon-based and machine learning approaches. Both methods view sentiment as either positive, neutral, or negative. According to the psycholinguistic approach, text sentiment is more continuous reflecting more naturally how we experience emotions. We compare these three approaches with a Twitter dataset collected during a service outage. Contrary to our expectation, we find that the language used in tweets is not very negative or emotionally intense. This research therefore contributes to the sentiment analysis discussion by dissecting three methods and discussing how and why they arrive at differing results. The selected research context provides an illuminating case about service failure and recovery.

Simone Griesser, Applied University of Northwestern Switzerland

Track 3: Machine Learning Methods
@SOPRA 2-3

A Machine Learning Technique to Classify LSST Observed Astronomical Objects Based on Photometric Data

The light curve analysis of the heavenly bodies is an indispensable tool for understanding the physical phenomena that govern them. Large telescopes like the LSST will produce an excess of data produced that will necessitate the need for automated methods to sift through it quickly and efficiently as doing so manually can be truly laborious. Keeping this in view, we have proposed an automated classification method using the simulated, photometric light curves in to 14 different classes. We have built our classification model by extracting several features and employing Random Forest classifier. Our proposed methodology performs reasonably well for most of the classes while others still offer a little room for improvement. As our proposed methodology relies on features extracted from photometric light curves, therefore it can be adapted and extended for use in other fields that rely on similar light curves.

Muhammad Usman Akram, National University of Sciences and Technology, Pakistan

Track 4: Data Science Operationalization
@BELLAVISTA 5

A Data Preparation Framework on Spark for Data Scientists

Traditionally, data scientists develop their models and the corresponding data preparation in interactive environments such as R or Python. These enable easy inspection of intermediate results and efficient testing of code chunks.

Production, on the other hand, places requirements such as stability, performance,

16

configurability and standardized logs. This is achievable by re-implementing the data pipeline for production in a high-performance language. For organizations with small analytics teams without a dedicated data engineering team, however, this procedure is too time-consuming and therefore not practicable.
In this talk, we will share our learnings on doing all the data preparation exclusively in Spark on Scala. In addition, we will present our Scala-based data preparation framework, which facilitates this process.
Kaspar Mösinger, Raiffeisen

Track 5: Data Ethics
@BELLAVISTA 2-3
Big Data Framing About Media Coverage in Switzerland and the USA
Big Data has become a renowned and commonly used term for describing large quantities of data and its collection for deeper analysis. The mass media play an important role in informing society about what Big Data is, as well as communicating the surrounding benefits and risks.
This paper presents an analysis of the media coverage of Big Data in Switzerland and the USA. A media content analysis, a frame analysis and a cluster analysis have been employed to examine the different frames within which the topic of Big Data has been presented in the Swiss and American Press as well as its evolution over time. The frame analysis shows that the majority of frames in both countries are not focusing on risks, but rather on benefits of Big Data.
Urs Dahinden, University of Applied Sciences HTW Chur

Track 6: Sponsored Talks
@ORIONE
GPU Acceleration with RAPIDS for Traditional Big Data Analytics or Traditional Machine Learning
The RAPIDS suite of open source software libraries gives you the freedom to execute end-to-end data science and analytics pipelines entirely on GPUs. It relies on It exposes that GPU parallelism and high-bandwidth memory speed through user-friendly Python interfaces. RAPIDS also focuses on common data preparation tasks for analytics and data science. This includes a familiar DataFrame API that integrates with a variety of machine learning algorithms for end-to-end pipeline accelerations without paying typical serialization costs. RAPIDS also includes support for multi-node, multi-GPU deployments, enabling vastly accelerated processing and training on much larger dataset sizes. The talk gives an overview about RAPIDS an how it can be used for machine learning.
René Müller, NVIDIA

Track 7: Sponsored Talks
@BELLAVISTA 4
AI computer vision with Microsoft Technology
Marc Schöni, Microsoft

Breakout Session 7 - 03:30pm - 04:00pm

Track 1: Machine Learning Applications
@ARENA

Fraud Detection in Financial Services Using Graph Analysis and Machine Learning

Both graph analysis and machine learning can be used very effectively to detect anomalies and outliers in datasets. The former is particularly useful when data can be represented as a network in which the connectedness of data, ie. the explicit relationships between entities, play a role. Networks of bank accounts connected by financial transactions are one obvious example, which is why modern fraud prevention applications use graph analytics and pattern matching on this kind of data. Graph analysis can be complemented by machine learning to yield results that are even more accurate.
In this session, we will show how a Spanish bank and an Eastern European provider of e-payments are using graph analysis for the purpose of fraud detection. We will look at how graph analysis and machine learning yield different results and how these techniques can be combined, provided the linked data can be turned into data structures that can be used in machine learning while at the same time maintaining the characteristics of the graph.

Hans Viehmann, Oracle

Track 2: Natural Language Processing
@SOPRA GRANDE

Harness the Power of User Generated Content with Search and Machine Learning

People share information digitally - within organizations (e.g. emails, CRM-entries, Wiki, Word/PDF, etc.), between customer and business or as user-generated content in the form of ratings and reviews. Unstructured text is everywhere, the volume is growing exponentially, but the resource is underutilized. One of the main reasons for that is that is the notorious difficulty of NLP methods. However, with the right set of (open-source) tools this has not to be true.
In this talk we will present a general workflow using search enhanced with ML-Methods to make insights in text based corpora accessible, extendable, and useful. We will describe how to set up a simple yet powerful NLP pipeline to ingest texts into ElasticSearch and explore Kibana for visualization. A powerful search cannot be underestimated when working with texts. Then we also will explain how to enrich the pipeline with ML-Methods (e.g. word embeddings, sentiment analysis, dependency parsing, network analysis, etc.). This process is supported by a specialized open-source library.
While we used this approach successfully in a proprietary product solution advisor tool we will demonstrate the approach by using review data from a popular website.

Jürgen Schwärzler, D ONE
Philipp Thomann, D ONE

Track 3: Machine Learning Methods
@SOPRA 2-3

Stress Pattern Recognition Through Wearable Biosensors in the Workplace: Experimental Longitudinal Study on the Role of Motion Intensity

Stress is a current issue in the workplace. Wearable biometric devices might improve stress monitoring, but it remains unclear how to recognize stress remotely. This longitudinal experimental study examined whether a non-physiological signal may improve the stress pattern recognition. The participants were 18 Public Administration employees using wearables for two months. The measures included established physiological signals (Galvanic Skin Response; Heart Rate) combined with a new non-physiological variable, associated with the user's physical activity (Motion Activity). Stress-related patterns were analyzed by adopting unsupervised learning approach with help of Gaussian Mixture Model and K-Means classification analysis, completed by the bootstrapping procedure. The results demonstrate that complementing physiological signals with physical activity-related information, improves stress pattern recognition, which may improve the quality of stress management data collection through Information Systems.

Vadym Mozgovoy, University of Lausanne

18

Track 4: Roundtables
@BELLAVISTA 5
Roundtables

- Table 1 - Can Data Analytics Improve Staffing and Planning?
  Alexander Grimm, Aspaara
- Table 2 - Data Science is done by humans. How to train data science teams effectively?
  Leo Marose, StackFuel
- Table 3 - When will the accountants be gone?
  Holger Rommel, ti&m
- Table 4 - How to attract and hire talent in today's candidate-driven job market?
  Melinda Braunke, Rockstar Recruiting
- Table 5 - Artificial Intelligence: Hype, Hope & Reality
  Matthias Brändle, Mobiliar

Track 5: Data Visualization
@BELLAVISTA 2-3
Generative Adversarial Networks: When Fake Never Looked So Real

In recent years a novel architecture has spurred increasing interest in the machine learning community. Generative Adversarial Networks have been used to tackle a plethora of different tasks: image synthesis from a text description, generation of super-resolution images, image compression, video slow motion; to name a few. In the years since GANs conception, the generation quality has reached unprecedented levels. This paves the way to alleviate one of the most important issues in real-life Machine Learning problems: lack of training examples.
In this talk, we will dissect GANs. We will talk about how they work, their shortcomings and their numerous applications. We will shortly discuss the implications of fabricated images in the era of fake news. Lastly, we will emphasize on how to utilize GANs in the industry.
Evangelos Ntavelis, CSEM

Track 6: Sponsored Talks
@ORIONE
Data Management and Why This Is So Strategic for Data Science Projects

Digitization, Big Data & Data Lakes, Data Driven Decision making are the buzzwords we hear every day. As part of the SAP Digital Platform, Data Management takes an important part and Michael will highlight in this session how this strategically fits into the SAP Standard Application landscape and why SAP is having such a large focus on this particular topic.
Additionally, the talk will be about showing the latest product development of the SAP technologies and how they can be used effectively in your non-SAP Data Science environment by talking about use cases and executed projects. A main focus of this talk will be how Data Science Projects can create value by integrating this into Business Processes.
Michael Probst, SAP

Track 7: Data Ethics
@BELLAVISTA 4
"All in" – The Data Ethics Poker Round

How to prepare your company for the ethical challenges of big data? Play your best cards in each round of the data life cycle to ensure that ethical standards are met. In this gamified discussion round, you will learn more about the new Ethical Codex for Data-Based Value Creation presented by the Swiss Data & Service Alliance. And yes: if you play the right cards, you can win!
Markus Christen, University of Zurich
Christoph Heitz, Zurich University of Applied Science

Breakout Session 8 - 04:00pm - 04:30pm

Track 1: Machine Learning Applications
@ARENA

Creating Value for Clients through Data & Analytics – How we Implemented Data-Driven Client Segmentation and Launched a Center of Excellence in Data & Analytics

Advanced analytics has increasingly been helping us understand client needs to support the way we serve our clients. While we've developed a strong level of expertise over the years, we found that pockets of skills and experience are not sufficient to capture the full potential of analytics. To further focus our efforts, we've launched a Center of Excellence in Data & Analytics. In this talk, we will present our journey towards a collaborative environment focused on creating value through analytics. As an illustrative example, we will discuss the typical use case of segmenting clients into groups with similar needs allowing us to offer the right service to every client. We will demonstrate our approach, discuss the impact and show how lessons learned helped us drive the design of the CoE.

Michel Neuhaus, UBS
Daniel Perruchoud, University of Applied Sciences Northwestern Switzerland

Track 2: Natural Language Processing
@SOPRA GRANDE

Syntax-Based Skill Extractor for Job Advertisements

In the context of extracting relevant skill-terms from job advertisements, we propose a syntax-based method for generating large amounts of machine-labeled text from a small amount of human-labeled data. This is then used to solve the vocabulary problem and significantly increase recall when detecting skills.

Ellery Smith, Zurich University of Applied Science
Thomas Haberthuer, Skillue

Track 3: Machine Learning Methods
@SOPRA 2-3

How to Optimize Gower Distance Weights for K-medoids Clustering Algorithm to Obtain Mobility Profiles of the Swiss Population

This piece of research aims to obtain mobility profiles of the Swiss population. To that end, a survey of the Swiss Statistical Office (FSO) called Mobility and Transport Micro-census (MTMC) is utilized. Along with a qualitative method clustering, the respondents in the survey are clustered based on their mobility characteristics to obtain their profiles. The clustering, in particular acquiring medoids (centrotypes or exemplars), helps us then to generate a synthetic population of Switzerland. To gain medoids of each cluster, the k-Medoids clustering algorithm is utilized which partitions instances based on their positions in a latent space (symmetric distance matrix). Distances that shape this space can be generated by various metrics e.g. Euclidean, Gower, Manhattan. Since in this study features are mixed-type (e.g. numeric, categorical, etc.), the Gower distance metric is preferred. In this study, the default weights of the Gower distance are optimized to obtain a higher Average Silhouette Width (ASW) value of the clustering results. ASW can be used to measure the quality of clustering results in which high value leads to higher intra-cluster homogeneity and inter-cluster dissimilarity. So, maximizing the ASW value improves the quality of the clusters which is the goal of the optimization. At the end, this process helps us to obtain more accurate mobility profiles of the Swiss population.

Alperen Bektas, University of Applied Sciences Western Switzerland

Track 5: Data Visualization
@BELLAVISTA 5

DNNViz: Training Evolution Visualization for Deep Neural Networks

In this paper, we present novel visualization strategies for inspecting, displaying, browsing, comparing, and visualizing deep neural networks (DNN) and their internal state during training. Despite their broad use across many fields of application, deep learning techniques are still often referred to as

"black boxes". Trying to get a better understanding of these models and how they work is a thriving field of research. To this end, we contribute with a visualization mechanism designed explicitly to enable simple and efficient introspection for deep neural networks. The mechanism processes, and displays neurons activation during the training of a neural network. We furthermore demonstrate the usefulness of this visualization technique through different use cases: class similarity detection, hints for network pruning and adversarial attack detection. We implemented this mechanism in an open source tool called DNNViz.

Gil Clavien, Fribourg University

Track 6: Sponsored Talks
@ORIONE

Towards Graph-based Machine Learning for Automated Health Care Services

We tackle the problem of predicting diagnoses for patients staying in critical care units. To this end, we employ healthcare data of 46 thousand patients with multiple admissions per patient. We consider multiple events as features per admission like fluids (e.g., insulin), lab tests (e.g., pH) and drugs (e.g., aspirin) which represent the evolving state of an admission. We employ graph data model to integrate additional information, per admission, from external sources (e.g., disease-symptom relations) and feed this enriched admission to a Recurrent Neural Network to predict diagnoses. Our approach shows significant results due to the relational information from previous admissions (connected via the external disease-symptom information).

Rhicheek Patra, Oracle Labs Zurich

Track 7: Data Ethics
@BELLAVISTA 4

"All in" – The Data Ethics Poker Round

How to prepare your company for the ethical challenges of big data? Play your best cards in each round of the data life cycle to ensure that ethical standards are met. In this gamified discussion round, you will learn more about the new Ethical Codex for Data-Based Value Creation presented by the Swiss Data & Service Alliance. And yes: if you play the right cards, you can win!

Markus Christen, University of Zurich
Christoph Heitz, Zurich University of Applied Science

Keynote Speaker & Award Ceremony - 04:30pm - 5:30pm

Keynote - Aleksandra Przegalinska

Context-Aware AI Systems of the Future

Technology is ubiquitous today. It can significantly increase human abilities and capabilities. This is not only an opportunity for positive change but also brings multiple issues and threats. Human-machine interaction research helps to address these challenges. Most important parts of it includes detecting physiological signals to uncover the uncanny valley effect as well as using deep learning for unravelling chatbot represented values integrity. The emerging concept of context awareness will be stressed upon. The future adaptive user interfaces will need to tailor a system which is responsive according to skills, tasks, and preferences of human users. A close cooperation between AI and IOT researchers will be indispensable for the future context-aware AI systems.

PRESENTING PARTNER

D ONE
www.d-one.ai

What comes first?

Is it the intelligent data scientists who are hungry for a challenge or the exciting problems that yearn for a solution?

Our clients come to us for the former and certainly do provide us enough of the latter. That's why our culture is focused on talent - because with the best people we can solve every problem. And it’s fun.

PLATINUM SPONSORS

Microsoft
www.microsoft.com

Microsoft

Microsoft Switzerland is a subsidiary of Microsoft Corporation (Redmond USA). Since its establishment in 1989, the subsidiary has developed into a typical Swiss SMB with 620 employees. Microsoft Switzerland is closely networked with the country 's economic and political institutions and maintains an active social dialogue on the topics of innovation, security and education. Microsoft is the world's leading manufacturer of standard software, services and solutions. We help people and companies from all industries and of all sizes to fully realize their potential.

Cloud, Data & AI Hands-on Labs
With the current pace of innovation on the Microsoft Azure platform, we want to make sure you and your team stay on top and continue leveraging what the platform has to offer. Therefore, we have set-up a wide range of workshops with the main and most frequently requested topic, Microsoft Azure.

These workshops are onsite trainings in Wallisellen or in Geneva and are delivered by the Swiss Technology Team or a Microsoft Partner. The aim of the Microsoft Cloud, Data & AI workshops is to expand knowledge of the Azure services through hands-on activities and develop an active Azure community.

You may visit www.aka.ms/ch-hands-on-labs and sign-up for one or more workshops. This website is getting renewed monthly with all updated and new events.
Our next workshop will be on April 26th, 2019 at Microsoft Vernier covering the topic Azure Fundamentals. The main objective of this workshop is that your teams get a solid understanding of Microsoft Azure that will enable faster value realization of Microsoft Azure platform for your organization.

PLATINUM SPONSORS

Migros-Genossenschafts-Bund
www.migros.ch

MIGROS

Data Strategy & Science
Our performance mandate is to generate insights from various internal and external data sources (transaction data, customer data, market data, online data ...) in order to support all business divisions of the Migros group in strategic, tactical and operational tasks. To fulfill this mandate, we develop automated and scalable analytics solutions based on machine learning to take advantage of all the opportunities coming along with the continuing digitization of Migros businesses.

Data Science
As specialists within the division Data Strategy & Science, data scientists take a wide variety of roles, depending on the project setup and the challenge at hand. From acting as a consultant or sparring partner in analytical projects to implementing and operating prototypes of new machine learning pipelines – we cover the whole range. In our daily work, we collaborate with other teams within our division (such as Web-Analytics) as well as colleagues from other departments at Migros (such as Marketing or Logistics). In addition, we connect with external partners in the private sector and academia to always remain on par with state of the art solutions.

Opportunities
The success of the Migros Group is based on the knowledge and skills of its employees. Women and men from 155 nations are committed every day to offering our customers products and services providing the best price-performance ratio. As the largest private employer in Switzerland, Migros also bears a special social responsibility which it fulfils with above average social benefits, secure workplaces and a working environment that is fair and characterized by respect.

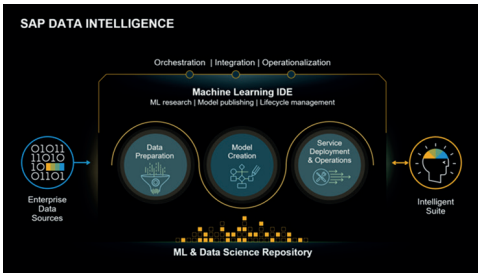
Find out more about the hundreds of open positions in more than 60 businesses of the Migros Group:
https://migros-gruppe.jobs

PLATINUM SPONSORS

SAP
www.sap.ch

From Data Management to AI with SAP Data Intelligence
If data is the new oil, it can be hard build your own drilling rigs, pipelines, refineries, and gas stations. Today, several point solutions for different aspects of the data management and AI lifecycle exist. But mixing and matching them can be a challenging experience.

We have created the **SAP Data Intelligence service** as the first enterprise solution with end-to-end coverage of the data management and AI life cycle. This next-generation offering aims to help customers and partners gain control of their data, models, and deployments. SAP Data Intelligence will consume data from all sources: SAP HANA, SAP HANA Cloud Services, SAP Business Warehouse, third-party databases, and data lakes based on Hadoop and the major cloud platform providers. Users will be able to prepare their data and build processing pipelines graphically and quickly, as well as deep dive into data science with notebooks. For this, SAP Data Intelligence will support Python, Scikit, Tensorflow, R, and an open container architecture. Model deployment, monitoring, and integration into the intelligent suite will be a breeze.



SAP Data Intelligence combines **SAP Data Hub** with **SAP Leonardo Machine Learning Foundation** in one next generation cloud service. Part of this comprehensive solution is the newly launched service called hierarchy matching, a machine learning service that can easily be used by non-data scientists to help with maintaining masterdata via automated predictions. SAP Data Intelligence is currently in beta, and is planned for general availability by end of June 2019. Like SAP Leonardo Machine Learning Foundation, it will be available via SAP Cloud Platform Enterprise Agreement. Find out how SAP Data Intelligence can help your organization gain intelligent insights and improve existing products through our Webinar series.

Learn more:
https://www.sap.com/swiss/products/leonardo/data-intelligence-monetization.html

#GiveDataPurpose

GOLD SPONSORS

Dataiku
www.dataiku.ch

Dataiku strives to bring large-scale adoption of analytics at scale to all enterprises through both self-service analytics and operationalization. In our experience with companies around the world, this approach is the difference that brings business-impacting change; mass adoption is more powerful than early adoption of specific new technologies.

Dataiku provides the technology that allows companies to:

- Dive in and get moving quickly on analytics, data science, machine learning, and AI at scale to stay competitive and overcome the fear of waiting for the right moment or the right technology.
- Effectively navigate and adapt to the perpetual stream of new technologies, allowing them to work with data efficiently at scale.
- Move quickly and iterate along the way, emphasizing and enabling putting good models in production (which is worth ten perfect models in the sandbox).

GOLD SPONSORS

ELCA
www.elca.ch

Since its founding in 1968, ELCA has continually grown. Today, it is one of the biggest independent Swiss full-service providers for business and technology solutions, and a leader in the fields of IT business consulting, software development and maintenance, and IT systems integration. The privately owned company, with more than 1000 experts, has branches in Lausanne, Zurich, Berne, Geneva, Paris, Madrid, Basel, Granada, Ho Chi Minh City and Mauritius (offshore development), all operating according to a common process framework.

ELCA develops innovative, custom-designed and high-performing business and technology solutions for a broad number of industries and customers. The solutions reduce complexity and costs, simplify the organization, and increase innovation cycles, improve business outcome and customer satisfaction. ELCA can develop individual software solutions as well as implement standard and open source technologies. More than 50 years of experience and a proven agile project and software development methodology enable ELCA to deliver projects successfully, including on a fixed-price basis, offering customers a high level of certainty in their budget and time planning.

ELCA puts great emphasis on identifying and recruiting top talents from leading Swiss and European universities. ELCA thoroughly trains new hires on its specific methodologies and approaches. ELCA's working methods are regularly reviewed with regard to quality assurance and the company is ISO 9001 and ISO 14001 certified. The company's development center in Vietnam, successfully in place for more than 20 years, has been under Swiss leadership from the beginning. The offshore branch has already attained Level 3 of the CMMI evaluation several times and offers ELCA and its clients the capacity to scale capabilities quickly and efficiently.

GOLD SPONSORS

NVIDIA

www.nvidia.com

GPU-ACCELERATE YOUR DATA ANALYTICS WORKFLOWS

Data science workflows have traditionally been slow and cumbersome when it comes to loading, filtering and manipulating data, as well as machine learning training itself. Explore GPU-accelerated machine learning and data analytics libraries, deployed on NVIDIA GPU, for maximized productivity, performance, and ROI.

RAPIDS: NEW SOFTWARE LIBRARIES FOR DATA SCIENCE

RAPIDS is built on more than 15 years of NVIDIA CUDA development and machine learning expertise. It's powerful new software for executing end-to-end data science training pipelines completely in the GPU, reducing training time from days to minutes.

THE WORLD'S FIRST PORTFOLIO OF PURPOSE-BUILT DEEP LEARNING SYSTEMS

Developed to meet the demands of AI and analytics, NVIDIA DGX Systems are designed to give data scientists the most powerful tools for AI exploration – from your desk to the data center to the cloud. Experiment faster, train larger models and arrive at insights – all starting on day one.

Reach out to our GPU-accelerated data science expert, Marc Stampfli mstampfli@nvidia.com, for more information, or find out more on

www.nvidia.com/datascience

GOLD SPONSORS

Oracle
www.oracle.com

Create tomorrow today
We have embedded innovative technologies into Oracle Cloud, enabling you to reimagine your businesses, processes, and experiences. Data management is at the core of any business. We care about the biggest asset: data. Our ambition is gaining more value from data while keeping it secure – on-premise and in the cloud. That is, why we developed a set of cloud services making you productive. The autonomous database is the newest and most advanced of them all. Oracle truly supports polyglot data management – you represent data as per your use-case. At SDS2019 we focus on our graph capabilities and use-cases.

Parallel Graph AnalytiX (PGX)
Graph analysis lets you reveal latent information that is encoded, not as fields in your data, but as relationships - metadata - between elements of your data - information that is not obvious to the naked eye, but can have tremendous value once uncovered.

What is PGX?
PGX is a toolkit for graph analysis - both running algorithms such as PageRank against graphs and performing SQL-like pattern-matching against graphs, using the results of algorithmic analysis. Graphs can be loaded from a variety of sources like flat files, SQL and NoSQL and Apache Spark and Hadoop; incremental updates are supported.

PATHWAYS TO INNOVATION
Oracle uniquely offers multiple paths and services to facilitate your move to the cloud. Integration, security, AI, ML, digital assistants and blockchain are key ingredients to our offering. Our applications will increasingly use AI suggesting the next best actions, automate answers, and provide personalized service.

Our cloud strategy is built on the understanding that cloud and on-premises implementations must coexist. This gives you a path that meets future needs while allowing you to preserve and upgrade your existing investments. Today, 430,000 customers in 175 countries use Oracle technologies to seize business opportunities and solve real, tangible challenges.

More information about
Cloud: www.oracle.com/cloud/
AI: www.oracle.com/artificial-intelligence/
PGX: www.oracle.com/technetwork/oracle-labs/parallel-graph-analytix/overview/index.html

GOLD SPONSORS

PwC

www.pwc.com

PwC's purpose is to build trust in society and solve important problems

We're a network of firms in 158 countries with more than 250,000 people who are committed to delivering quality in assurance, advisory, tax & legal and digital services. Within PwC Switzerland more than 3,200 employees and partners in 14 locations in Switzerland and one in the Principality of Liechtenstein help to create the value organizations and individuals are looking for.

PwC Digital Services in Switzerland is part of the worldwide community of experts dedicated to help you in the face of digital disruption and 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.

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

- **Innovation Lab**: helping organizations accelerate their data analytics innovation
- **Business Solutions**: helping organizations solve business problems with our data analytics applications
- **Strategy through Execution**: helping organizations 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 its implementation. In Switzerland we're PwC's centre of excellence for machine learning, artificial intelligence, optimization, simulation and modeling.

Reach out to our team of data science and analytics experts:

Christian Westermann, Partner and Leader Data and analytics, +41 58 792 27 97
christian.westermann@ch.pwc.com

Christian Blakely, Senior Manager, AI and real-time analytics, +41 58 792 22 81
christian.blakely@ch.pwc.com

GOLD SPONSORS

Siemens
www.siemens.com

SIEMENS
Ingenuity for life

Siemens Mobility is a separately managed company of Siemens AG. As a leader in transport solutions for more than 160 years, Siemens Mobility is constantly innovating its portfolio in its core areas of rolling stock, rail automation and electrification, turnkey systems, intelligent traffic systems as well as related services. With digitalization, Siemens Mobility is enabling mobility operators worldwide to make infrastructure intelligent, increase value sustainably over the entire lifecycle, enhance passenger experience and guarantee availability. In fiscal year 2018, which ended on September 30, 2018, the former Siemens Mobility Division posted revenue of €8.8 billion and had around 34,200 employees worldwide.

Make mobility safer, faster, more convenient – and fun

Today, people are looking for solutions for everyday mobility that are simple, flexible, fast, reliable and affordable. At the same time, countries and cities are facing the challenge of reducing cost pressure, space requirements, traffic-related noise and CO2 emissions.

The pressure on mobility providers and decision-makers to meet these mobility and transportation needs is not only high but also increasing – by 2050 the urban population is expected to exceed 70%. Facing these rising demands, the transport industry is now looking for solutions that will take existing transport infrastructure to the next level.

Digitalization not only plays a key role for mobility providers within their struggle to achieve the reliability, flexibility and availability they need to make rail solutions cost-efficient: It is also the key to make traveling easier and more pleasant for passengers. Locomotives, trains, components, train automation systems – They are all valuable data sources.

With Siemens Digital Services, a state-of-the-art system that processes this wealth of transportation data, transport operators analyze trends and monitor the status of machinery and the flow of passengers and goods to provide a smooth mobility experience.

Further information is available at: www.siemens.ch/mobility

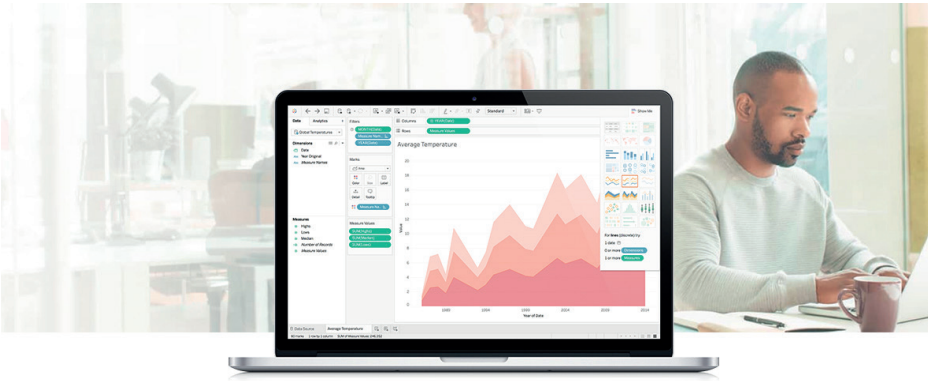
GOLD SPONSORS

Tableau
www.tableau.com

Tableau helps people and organizations become more data-driven as the trusted leader in analytics. The Tableau platform provides the breadth and depth of capabilities to serve the needs of even the largest global enterprises in a seamless, integrated experience. Tableau is designed to fit, not dictate your data strategy, and adapts to your environment with unmatched flexibility and choice, while meeting the toughest governance and security requirements. People love using Tableau because it is both powerful and intuitive—and offers a fundamentally different user experience by empowering people of all skill levels to explore and analyze data using visuals and natural language. Tableau has become the standard language of analytics for modern business users and continues to lead the industry with the most passionate and engaged user community in analytics, a customer base with millions of users at more than 86,000 organizations, and a deep commitment to customer-focused innovation.

For more details please visit:

https://www.tableau.com



SILVER SPONSORS

la Mobilière
www.mobiliere.ch

la Mobilière: Live your life. We'll be here.

La Mobilière is the most personal insurance company in Switzerland. Founded in 1826, la Mobilière is Switzerland's first private insurance provider and has over 190 years of expertise. The company has grown to the Swiss market leader in several major insurance segments. It is organised as a cooperative and customers directly participate in its success. With 79 general agencies at 160 locations throughout Switzerland, la Mobilière serves two million customers and nine out of ten claims are handled locally.
The Data & Analytics team at la Mobilière establishes and enables business-driven analytics as part of the company wide digital transformation. In order to create value from data, we identify and work on data-based use cases combining state of the art technologies with the business's long standing insurance expertise. With the benefits for our customers in focus, we optimise and innovate processes and support decision-making.

NZZ
www.nzz.ch

Since 1780, the "Neue Zürcher Zeitung" has stood for well-founded, precise and detailed reporting on international, national and regional events. In collaboration with a worldwide network of correspondents, the "Neue Zürcher Zeitung" offers daily authentic reports, clever analyses, clear commentaries and exciting reports. The "Neue Zürcher Zeitung" is one of the most renowned national and international voices in quality journalism.

SILVER SPONSORS

Raiffeisen
www.raiffeisen.ch

RAIFFEISEN

Raiffeisen: third-largest banking group in Switzerland

The third largest force in the Swiss banking sector services roughly half of the Swiss population. 1.9 million Cooperative members are co-owner of their Raiffeisen bank. The Raiffeisen Group is present at 880 locations throughout Switzerland. The 246 legally autonomous cooperative Raiffeisen banks are amalgamated into Raiffeisen Switzerland Cooperative, which is the strategic leader of the entire Raiffeisen Group. Through group companies, partnerships and participations, Raiffeisen offers private individuals and corporate clients the full range of products and services. As of 31.12.2018, the Raiffeisen Group had CHF 196 billion in assets under management and CHF 188 billion in loans to clients. The market share in the mortgage business amounts to 17.6 percent. Total assets are CHF 225 billion.

SAS
www.sas.ch

SAS is the global market leader in analytics and one of the largest software manufacturers with sales of USD 3.24 billion. Customers worldwide use innovative software and services from SAS to transform data into knowledge and make intelligent business decisions. Since 1976 SAS has provided customers around the globe with THE POWER TO KNOW.

With SAS, companies develop and implement strategies, measure their own success, make their customer and supplier relationships profitable, manage the entire organization in real time and meet regulatory requirements. The US parent company is based in Cary, North Carolina. SAS Switzerland has its headquarters in Zurich and another branch in Geneva. More info on www.sas.com

At this year's SDS conference, SAS introduces the SCYP – SAS Certified Young Professionals – Program.

The demand for individuals with SAS knowledge and expertise is growing in organizations globally. To help close the analytics skills gap, SAS is offering students access to software, training and certification - for free. Enrich your analytical skills by becoming a SAS professional and maximize your potential in the job market. Register now: www.sas.com/scyp

SILVER SPONSORS

SBB
www.sbb.ch

Mobility is undergoing a profound and comprehensive change. New customer needs, digitalisation, new mobility providers and regulatory developments are changing markets and business models. The SBB 2020 strategy and our vision provide some answers to this. We shape the mobility of the future – simple, personal, connected.

SBB is the backbone of the Swiss public transport system, and day-to-day rail operations are the basis of what we do. SBB has been transporting people and freight for more than 100 years. By doing so, we are making an important contribution to the quality of life and competitiveness in Switzerland. We want to continue this success story, even in times when the entire economy and society, including the mobility industry, are undergoing profound changes. We will therefore continue to fulfil our responsibility towards public transport and Switzerland in future.

Skillue
www.skillue.com

Unleash your potential using an AI powered skill-to-job matching engine
Skillue uses data science technologies, like Artificial Intelligence and machine learning, to create and develop a digital marketplace that enables talent, job and enterprise benchmarking based on demanded and offered skills. As a result, Skillue is building a new meeting zone for candidates and companies.

Create your career path
Tell us what you want, and we will tell you what to do. It's as simple as that. Based on your career path options, we will guide you to the skills you need to be 100% fit for the jobs you want.

Benchmark yourself to the market
Try our AI-powered benchmarking utility to see where you stand, compared to peers on the market. Then, just go ahead and improve the skills you need in order to get the job you want.

100% Anonymity
And all that anonymously...
Decide when to share and who to share your full CV. You can verify exactly what companies can see your profile. We offer just enough information to our clients about your profile to get them interested. The rest is up to you. You can choose to share your personal details and CV with employers at any time.

SILVER SPONSORS

SwissRe
www.swissre.com

Swiss Re

The Swiss Re Group is one of the world's leading providers of reinsurance, insurance and other forms of insurance-based risk transfer, working to make the world more resilient. It anticipates and manages risk – from natural catastrophes to climate change, from ageing populations to cyber crime. The aim of the Swiss Re Group is to enable society to thrive and progress, creating new opportunities and solutions for its clients. Headquartered in Zurich, Switzerland, where it was founded in 1863, the Swiss Re Group operates through a network of around 80 offices globally. It is organised into three Business Units, each with a distinct strategy and set of objectives contributing to the Group's overall mission.

ti&m
www.ti8m.com

ti&m
big ideas. creative technology.

We are market leaders in digitisation and security products as well as in innovation projects in Switzerland and the financial centres of the EU. For our demanding customers, we vertically integrate the entire IT value chain. In our offices in Zurich, Bern and Frankfurt am Main, we currently employ over 320 outstanding engineers, designers and consultants. And further offices in Europe will follow. Our growth is based on our strengths and values: courage, creativity, agility and entrepreneurial flair coupled with sustainability and Swissness.

SILVER SPONSORS

Wincasa
www.wincasa.ch

Innovative solutions spanning the real estate life cycle
Wincasa is the leading integrated real estate service provider in Switzerland. With 860 specialists, we offer customers an extensive service portfolio covering the entire lifecycle of properties, from initial planning, construction and management to revitalisation and repositioning. 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. The company has been increasingly focusing on digital leaders from other sectors, systematically pursuing digitalisation 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.

COMMUNITY SPONSORS

Euresearch
www.euresearch.ch

EURESEARCH
Swiss guide to European research & innovation

Euresearch – your Swiss guide to European research and innovation

Euresearch is an information and advisory service on the European Research and Innovation Framework Programmes supported by the Swiss federal government. Euresearch has Offices all over Switzerland, including advisors with specific services for companies and a Network Office in Bern.

Horizon 2020 is the biggest EU Research and Innovation programme with nearly €80 billion of funding over 2014 to 2020 (and €14 billion for 2020). Numerous opportunities for Information and Communication Technologies (ICT) innovation and their applications are available for Swiss participants from the public and private sector.

Benefit for you?: if you are a researcher or innovator in Switzerland you can request specific information on European funding opportunities free of charge.

Get in touch:
Contact: info@euresearch.ch

SATW
www.satw.ch

satw it's all about technology

The Swiss Academy of Engineering Sciences SATW is the most important network of experts for engineering sciences in Switzerland and is in contact with the highest Swiss bodies for science, politics and industry. The network is comprised of selected individual members, member organisations and experts.

On behalf of the federation, SATW identifies industrially relevant technological developments and informs politics and society about their importance and consequences. As a unique expert organisation with high credibility, it conveys independent and objective information on technology – as the basis for establishing well-founded opinions. SATW also promotes the interests and understanding of technology in the population, including young people in particular. It is politically independent and non-commercial.

COMMUNITY SPONSORS

SGAICO
https://sgaico.swissinformatics.org

swiss group for artificial intelligence and cognitive science

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 AI in Switzerland to discuss and experience the technology everybody talks about!

Contact: sgaico@swissinformatics.org

SKDV
www.skdv.ch

SKDV
schweizer kundendienst verband

Your successful customer service is our mission!
The Swiss Customer Service Association SKDV promotes and supports its members in innovation and optimization around customer service in Switzerland.

Successful customer service in the digital age is more than process optimization and automation. Successful customer service is a lived attitude in all functions of your company. The seamless integration of these factors becomes the benchmark for the quality of your service.

The SKDV supports you as a specialist and manager in the service area in the design of your customer service in all industries.

The SKDV strengthens the position of the service sector in Switzerland - join us now!
www.skdv.ch

COMMUNITY SPONSORS

SSS

www.stat.ch

SWISS STATISTICAL SOCIETY
www.stat.ch

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 / industries).
- organises short courses oriented towards practical applications.
- supports a yearly seminar of PhD students.

ORGANIZER

Swiss Alliance for Data-Intensive Services
www.data-service-alliance.ch

The Swiss Alliance for Data-Intensive Services is a technology network for innovative companies, academic institutes and individuals with a focus on data-driven value creation: services, products and business models based on digital data. It is a community that helps companies to move forward with digitalization and brings key innovators together.

In doing so, we rely on three pillars for our success:

- R&D projects by our members for pushing forward innovation, and cooperation within an interdisciplinary network of experts from innovative companies and universities to combine knowledge from different fields into marketable products and services.
- Top employees and best-in-class education.
- Inspiration and exchange via connecting domain experts and joint workshops, conferences and Expert Groups such as Machine Learning Clinic, Data Ethics, Natural Language Processing and Predictive Maintenance.

To boost innovation several initiatives are launched in 2018, e.g., the initiative “From ideas to projects” and a start-up grant 2018/2019. The Swiss Alliance for Data-Intensive Services makes a significant contribution in creating data-driven added value in Switzerland.

Get in touch:
More information: data-service-alliance.ch

Contact: Gundula Heinatz Bürki (gundula.heinatz@data-service-alliance.ch)

Organization Committee

Melanie Geiger, Gundula Heinatz, Amrita Prasad

SDS2020

We hope you enjoyed SDS2019 and got useful insights and contacts. If you have ideas for participation or comments on the conference, please send them to

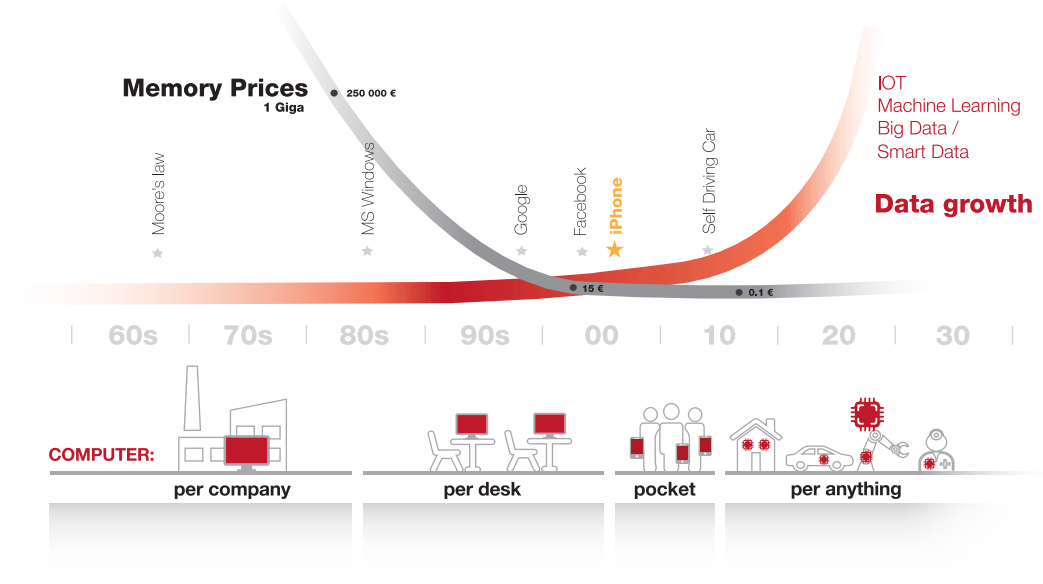
conference@data-service-alliance.ch

Thank you!

Imprint

Swiss Alliance for Data-Intensive Services
Bälliz 62
3600 Thun
Phone: +41 33 221 88 20
Email: gundula.heinatz@data-service-alliance.ch

DATA DRIVES DIGITIZATION



D|ONE

WE MAKE SENSE.

DATA DRIVEN VALUE CREATION

DATA SCIENCE & ANALYTICS | DATA MANAGEMENT | VISUALIZATION & DATA EXPERIENCE

D ONE, Sihlfeldstrasse 58, 8003 Zürich, d-one.ai