

SwissText 2019

Fourth Swiss Text Analytics Conference

Program

June 18 & 19, 2019

Content

Preface	3
Schedule	4
Keynotes	6
Presentations	10
Posters	25
Demos	28
Sponsors	32
Partners	37
Feedback	48
Impressum	49
Program Committee	50

Preface

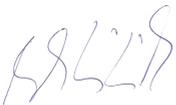
Welcome to the 4th Swiss Text Analytics Conference!

The field of Text Analytics and Natural Language Processing is rapidly progressing and adapting. Four years ago, when we organized the first SwissText conference, we called it “Swiss Text Analytics Conference”, since much of the research and applications were focussed on “Text Analytics” at the time: text classification, sentiment analysis and named entity recognition were the predominant topics in the 2016 edition. Since then, every year topics have shifted towards more complex and challenging areas. This year, there seems to be a major interest in speech processing, especially for standard German and Swiss German, where we will have eight presentations and a workshop.

The program of this year's conference is again crammed with interesting events and highlights: there will be more than 40 presentations and posters from industry and research, and three high-profile keynotes; we will run a Demo Session, in which more than 15 companies and research institutions will present their solutions live and interactively; we will have workshops on NLP for Business and on Swiss German, as well as a tutorial for building chatbots; and, finally, we will host the evaluation campaign and award ceremony of the shared task for Text Summarization in German.

Organizing all these activities involves a huge amount of effort, and I would like to thank all contributors for their great support: our sponsors and partners, ZHAW for hosting us, the reviewers, the local organizing committee, and of course the speakers and presenters!

I am now looking forward to an exciting conference with fascinating presentations and lots of interesting discussions!



Mark Cieliebak

Conference Chair



Schedule

Tuesday - Day One, June 18, 2019

08:30	Registration & Coffee/Gipfeli - Foyer		
09:00	Welcome Message: Mark Cieliebak (ZHAW) - Room TN E0.46 - E0.54		
09:30	Keynote: Lisa Gibbs (The Associated Press) - Room TN E0.46 - E0.54		
10:15	Break		
10:45	Parallel Sessions		
	Track 1 - Room TN E0.46	Track 2 - Room TN E0.54	Track 3 - Room TN E0.58
	<i>Didier Orel, Titus Plattner and Marcel Blattner:</i> NLP Activities at Tamedia	<i>Timo Spring, Jacky Casas, Karl Daher, Elena Mugellini and Omar Abou Khaled:</i> Empathic Response Generation in Chatbots	<i>Tatyana Ruzsics, Massimo Lusetti, Anne Göhring, Tanja Samardzic and Elisabeth Stark:</i> Neural Text Normalization with Adapted Decoding and POS Features
	<i>Ahmad Aghaebrahimian:</i> Hyperparameter Tuning for Deep Learning	<i>Christoph Süess:</i> Chatbots usability - Dos and don'ts	<i>Christof Traber, Schamai Safra, Bleicke Holm, Dominic Schnyder and Philipp Lichtenberg:</i> Text-to-Speech (TTS) for Seven Swiss German Dialects
	<i>Nils Schaetti:</i> Author Verification in Stream of Text with Echo State Network-based Recurrent Neural Models	<i>Laura Gander, Antonella Bolt and Ursula Stäubli:</i> Customer-driven architecture for a Chatbot	<i>Stephan Wick and Samuel Läubli:</i> Neural Machine Translation for Increased Human Translation Efficiency at Migros Bank
12:15	Lunch Break		
13:30	Poster Presentations - Room TN E0.46 - E0.54		
14:00	Poster Session - Foyer		
15:00	Coffee Break		
15:30	Workshops		
	Room: TN E0.46	Room: TN E0.54	Room: TN E0.58
	1st Business NLP Workshop Organizer: <i>Swiss Alliance for Data-Intensive Services</i>	Joint Actions for Swiss German Organizers: <i>UZH, ZHAW</i>	Building A Simple Chatbot With Dialogflow Organizers: <i>SpinningBytes AG, ZHAW</i>
4 17:00	End of Day One		

Schedule

Wednesday - Day Two, June 19, 2019

08:30	Registration & Coffee/Gipfeli - Foyer		
09:00	Welcome Message: Mark Cieliebak (ZHAW) - Room TN E0.46 - E0.54		
09:30	Keynote: Hinrich Schütze (LMU) - Room TN E0.46 - E0.54		
10:15	Break		
10:45	Parallel Sessions & Shared Task on German Summarization		
	Track 1 - Room TN E0.46	Track 2 - Room TN E0.54	Room TN E0.58
	<i>Deepanwita Datta, Manajit Chakraborty and Aveek Biswas:</i> Your Click Matters: Enhancing Click-based Image Retrieval performance through Collaborative Filtering	<i>F. Rinaldi, P. Kuntschik, G. Jürgen, M. Leddin, R. Rodriguez Esteban, Al. Weichselbraun, T. Ellendorff, N. Colic and L. Furrer:</i> MedMon: social media analytics for a healthcare application	Shared Task on German Text Summarization
	<i>Aaron Richiger:</i> Automated Data Extraction from Documents for Analysis using Machine Learning	<i>Gion Linder:</i> Success of ASR (automatic speech recognition) in the field of subtitling	Organizers: <i>Dominik Frefel, Lukas Neukom, Manfred Vogel FHNW</i>
	<i>Gabriel Luthier and Andrei Popescu-Belis:</i> PLACAT: A user-friendly question answering system for smart speaker devices	<i>Andrew Marritt:</i> Building an inductive coding service for employee feedback	
12:15	Lunch Break		
13:30	Special Session: Legal Issues with Data - Room TN E0.46 - E0.54		
14:00	Demo Presentations - Foyer		
15:30	Coffee Break		
16:00	Keynote: Mona Diab (GWU/Amazon) - Room TN E0.46 - E0.54		
16:45	Closing Statement - Room TN E0.46 - E0.54		
17:15	Apéro		
18:45	End		

Keynote

Lisa Gibbs

More News, More Efficiently: Text Automation at the Associated Press

Abstract:

The Associated Press has been producing stories using text automation since 2014. While the news agency remains focused on corporate earnings and sports stories, it's moving into some new content areas – such as how automation and AI might help U.S. local news outlets cover their communities better. And, it's perfecting a news summarization tool that automatically generates summaries of every AP text story. Along the way, AP has learned much about assessing the viability for an automation project, the challenges of data management, the ethics of news automation and more.

Biography:

Lisa Gibbs is Director of News Partnerships at The Associated Press and the newsroom's point person on AP's automation and artificial intelligence strategy group. As AP's business editor, she oversaw AP's first significant text automation initiative from 2014, which uses automation software to generate roughly 3,700 corporate earnings stories every quarter. Current projects span a range of technologies, including image recognition, machine learning for verification of social media and event detection. Gibbs is a frequent speaker on how newsrooms can use AI technologies to produce better journalism, more efficiently. Recent talks include sessions at South by Southwest, the International Symposium on Online Journalism, Computation + Journalism Symposium and the Google News Initiative Innovation Forum.

Organization:

The Associated Press

Contact:

lgibbs@ap.org



Keynote

Hinrich Schütze

Byte/Character/Subword/Morpheme/Word/Phrase: Which Representational Unit is Best for Deep Learning?

Abstract:

At the lowest level of Deep Learning architectures for Natural Language Processing, language is usually represented by an embedding layer. Embedding units of different granularity are in use. Recent advances in neural machine translation seem to suggest that characters perform best. But character-level models are difficult to train and interpret, and there is no strong evidence that they work well for semantic tasks such as natural language inference. After an analysis of the pros and cons of different granularities from the perspectives of linguistics and machine learning in the first part of the talk, I will then present an experimental study of BERT, a contextualized embedding model whose embeddings are wordpieces, units of a granularity between characters and words. I will discuss problems we have found with wordpieces and show how they can be addressed.

Biography:

Hinrich Schütze is Professor of Computational Linguistics and Director of the Center for Information and Language Processing at LMU Munich in Germany. Before moving to Munich in 2013, he taught at the University of Stuttgart. He received his PhD in Computational Linguistics from Stanford University in 1995 and worked on Natural Language Processing and information retrieval technology at Xerox PARC, at several Silicon Valley start-ups and at Google (1995-2004 and 2008/9). He is a co-author of two well-known books: "Foundations of Statistical Natural Language Processing" (with Chris Manning) and "Introduction to Information Retrieval" (with Chris Manning and Prabhakar Raghavan).



Organization:

Ludwig-Maximilians-Universität München

Contact:

<https://cis.lmu.de/schuetze/contact.html>

Keynote

Mona Diab

Low Resource Scenarios: Challenges and Opportunities

Abstract:

With the advent of social media, we are witnessing an exponential growth in unstructured data online. A huge amount of this data is in fact in languages other than English. Some of these languages have rich automated resources and processing tools, but the majority of the languages in the world are considered low resource despite presence online. In this talk, I will address the problem of processing low resource languages. I will present some of our solutions for language identification, information extraction, machine translation, and resource creation exploiting rich languages via cross-linguistic modeling. Such techniques can also be cast for cross-genre and cross-domain challenges.

Biography:

Mona Diab is Principal Scientist at Amazon AWS AI Deep Learning in Seattle, WA, USA. She is also Professor of Computational Linguistics and NLP in the Department of Computer Science, George Washington University (GW), and founder and Director of the NLP lab CARE4Lang at George Washington University. Before joining GW, she was Research Scientist (Principal Investigator) at the Center for Computational Learning Systems (CCLS), Columbia University in New York. She is also co-founder of the CADIM group with Nizar Habash and Owen Rambow, one of the leading reference points on computational processing of Arabic and its dialects. Her research interests span several areas in computational linguistics/natural language processing: cross-linguistic modeling and multilingual processing, computational lexical semantics, social media processing, information extraction & text analytics, machine translation, resource building, and computational socio-pragmatics. She has a special interest in low-resource language processing with a focus on Arabic dialects. Mona has more than 200 scientific publications and she serves as acting editor for several of the community's top scientific journals such as TACL and CSL.

Organization:

Amazon / George Washington University

Contact:

mtdiab@gwu.edu





SUBTITLING AND MORE **ACCESS SERVICES AT SWISS TXT**

- People with a sensory handicap want to have access to multimedia content. Without access services such as subtitles, audiodescription and sign language this would not be possible.
- SWISS TXT has been providing access services for the public broadcaster SRG SSR for 30 years.
- We have around 100 employees at our regional locations (Zurich, Geneva, Comano) so we can provide professional services in Switzerland's 3 official languages and around 20 other languages.



SWISS TXT subtitles about 65% of the public broadcaster SRG SSR. That is about 37'000 hours a year.



Will ASR (automatic speech recognition) and AI-supported workflows in the field of subtitling be the future?



YOUR CONTACT **GION LINDER**

Head of Access Services
T: +41 32 329 22 02
E: gion.linder@swisstxt.ch

SWISS TXT

SWISS TXT AG
Alexander Schöni-Strasse 40
Postfach 1136
CH-2501 Biel-Bienne
T +41 32 329 29 29
www.swisstxt.ch

Presentation

Didier Orel, Titus Plattner, and Marcel Blattner

Automatic News Generation & Reader's Comments Processing

Abstract:

In this talk, we will survey how Tamedia plans to benefit from recent technologies for Natural Language Generation to implement new and innovative methods for news generation. Tamedia is the largest private media group in Switzerland, owning more than 30 newspapers in the country: Tages Anzeiger, 20 Minuten, Sonntagszeitung, Tribune de Genève, Berner Zeitung, Basler Zeitung, 24 heures,

Tobi is Tamedia's automated text project for vote results at a local level. On November 25th 2018, in just five minutes, we generated about 40.000 texts – with customized variations for all 2222 municipalities in two languages. Three months later, we did the same, including even cantonal and local results.

These successful pilots for vote results demonstrated how we can address a long tail audience with localized news. These texts were generated with a template-based approach. Other media organizations, with Associated Press at the vanguard since 2014, already use this approach. But we think that a personalized distribution of these texts – with for example personalized push alerts or a tailored feed – will allow leveraging a large quantity of micro-audiences. The key success factors of these pilots were: the experienced journalists who crafted the templates; smooth user experience thanks to an elegant and clear front-end; and the measure of user feedback. Based on the positive outcome of the pilots, Tamedia plans to automate texts for sports results.

Tamedia is also enabling advanced models of text analysis to process the huge amount of comments readers write on the different digital channels. Goals are on one hand to detect toxic comments in order to support the moderation process and speed up the review time, and on the other hand to surface the most interesting comments and extract a snippet from them (see <https://twitter.com/tagicomments>).

Biography: Didier Orel is project and product manager at Tamedia, leading private media group in Switzerland. He's focusing on innovation projects for Media, leveraging data and machine learning potentials in media industry. Among the topics launched in this field are: automatic generation of contents, image and text search engines and news recommendation/personalization engines.

Titus Plattner is a reporter and innovation project manager at Tamedia. In 2017-2018, he was awarded with a one-year JSK fellowship at Stanford University, where he focused on the interface between journalism and computer science with project related to personalization and automatic content creation, among others.

Dr. Marcel Blattner studied physics and mathematics at the University of Zurich. As Chief Data Scientist at Tamedia Digital, Marcel Blattner is currently responsible for the development and implementation of complex data analyses that generate direct added value for all stakeholders. He is author of several publications (<http://bit.ly/1zzo1MD>)

Organization: Tamedia

Contact: didier.orel@tamedia.ch, titus.plattner@tamedia.ch,
marcel.blattner@tamedia.ch

Presentation

**Timo Spring, Jacky Casas, Karl Daher, Elena Mugellini,
and Omar Abou Khaled**

Empathic Response Generation in Chatbots

Abstract:

Recent years show an increasing popularity of chatbots, with latest efforts aiming to make them more empathic and human-like, finding application for example in customer service or in treating mental illnesses. Thereby, empathic chatbots can understand the user's emotional state and respond to it on an appropriate emotional level. This survey provides an overview of existing approaches used for emotion detection and empathic response generation. These approaches raise at least one of the following profound challenges: the lack of quality training data, balancing emotion and content level information, considering the full end-to-end experience and modelling emotions throughout conversations. Furthermore, only few approaches actually cover response generation. We state that these approaches are not yet empathic in that they either mirror the user's emotional state or leave it up to the user to decide the emotion category of the response. Empathic response generation should select appropriate emotional responses more dynamically and express them accordingly, for example using emojis.

Biography: Jacky Casas and Karl Daher are pursuing PhDs at the University of Fribourg in collaboration with HES-SO. Jacky's research interests are AI, NLP and chatbots while Karl's main interest is empathy between machines and humans. This survey was done by Timo Spring, Master student in computer science at University of Bern, and is a collaboration between the two worlds: empathy and chatbots.

Organization: University of Bern and HES-SO, University of Applied Sciences Western Switzerland

Contact: timo.spring@students.unibe.ch, jacky.casas@hes-so.ch,
karl.daher@hes-so.ch, elena.mugellini@hes-so.ch,
omar.aboukhaled@hes-so.ch

Presentation

Tatyana Ruzsics, Massimo Lusetti, Anne Göhring, Tanja Samardzic, and Elisabeth Stark

Neural Text Normalization with Adapted Decoding and POS Features

Abstract:

The task of text normalization aims at bringing non-canonical language, coming from speech and social media, to a standardized writing. This task is especially important for languages such as Swiss German, with strong regional variation and no written standard.

In this work, we propose a novel solution for normalizing Swiss German WhatsApp messages using the encoder-decoder neural machine translation (NMT) framework. We enhance the performance of a plain character-level NMT model with the integration of a word-level language model and linguistic features (POS tags). The two components are intended to improve the performance by addressing two specific issues. The former targets the fluency of the sequences predicted by NMT: it corrects a sequence which is not a proper word, despite being a likely sequence of characters. In addition, this modification targets the frequent cases where a contracted form corresponds to multiple normalized words, e.g. the word 'kömmer' ('we can') is mapped to the normalization form 'können wir'. The latter component, the addition of POS tags, aims at resolving cases of word-level ambiguity. For example, the ambiguous input word Lüüt can be normalized as the noun Leute 'people' or läuten 'to ring' when used as a verb. Our systematic comparison shows that the proposed model improves over the best previous solution. A thorough analysis of the compared systems' output shows that our two components produce indeed the intended, complementary improvements.

Intended audience:

Our intended audience includes both researchers and professionals interested in applying our resources in their own work.

Biography: Massimo Lusetti obtained a Master's degree in Multilingual Text Analysis at the University of Zurich, where he currently works as assistant. His main interests in NLP are machine translation and distributional semantics.

Organization: University of Zurich

Contact: tatiana.ruzsics@uzh.ch, massimo.lusetti@uzh.ch, goehring@cl.uzh.ch, tanja.samardzic@uzh.ch, estark@rom.uzh.ch

Presentation

Ahmad Aghaebrahimian

Hyperparameter Tuning for Deep Learning in Natural Language Processing

Abstract:

Although Deep Learning has advanced a lot through past several years, it still seems like a black art for many people mostly due to the fact that obtaining consistent and good results from a deep architecture requires optimizing many parameters. Hyperparameter tuning is an essential task in deep learning which can lead to significant changes in the network performance. This talk is the essence of approximately 3000 GPU hours on optimizing a network for a text classification task on a wide array of hyperparameters. Word embedding types, word embedding sizes, word embedding updating, character embedding, character embedding sizes, deep architectures (CNN, LSTM, GRU), optimizers, gradient control, classifiers (Softmax, Sigmoid, CRF), dropout, deep vs. wide networks, pooling, and batch sizes are the hyperparameters studied in this work using a grid search scheme. I will talk about the most critical parameters and the insight behind them that researchers can modify or prioritize in a deep architecture to get the best performance with the least effort on the part of humans and the machine.

Biography: Dr. Ahmad Aghaebrahimian is a Research Associate at the Zurich University of Applied Sciences (ZHAW) with backgrounds in computer science and linguistic. He completed his Ph.D. in Mathematical Linguistics at the Charles University in Prague where he integrated Deep Neural Networks and linguistics to improve the state of the art in large-scale open-domain Question Answering. His areas of interest include Artificial Intelligence in general and Deep Neural Networks in particular, as well as Corpus Linguistics, Question Answering, and Information Retrieval.

Organization: Zurich University of Applied Sciences

Contact: agha@zhaw.ch

Presentation

Christoph Süess

Chatbots Usability - Dos and Don'ts

Abstract:

More than twenty-five years ago the first web pages popped up in the internet. The process to deploy a web site was rather technical and topics like usability and user experience did not get a lot of attention. This changed a lot. Nowadays is all about user-centered design.

With chatbots the history repeats itself. The first chatbots were created by programmers who were fascinated by the technical possibilities. But how can we develop user-friendly chatbots?

In this presentation Christoph Süess will explore popular Swiss chatbots related to different industries such as banking, insurance, tourism and mobility. Join the talk, take a look at chatbot conversations and experience smart and less clever practices. Learn how wording, dialogue structure, fallback strategies, the use of quick answers and other techniques can impact the user experience.

This talk will arm you with a bag of tips and tricks to build chatbots which boost the user engagement, which appear intelligent and trustworthy and that help you reach your business goals.

Biography: Christoph Süess is a software engineer and the cofounder of the swiss based software company Paixon. He completed his studies in computer science at the university of applied sciences of Rapperswil in 2011. After several years as a software engineer he co-founded Paixon. Paixon develops custom softwares such as web applications and mobile apps. Since 2017 the company implements chatbots for various Swiss companies and shares its knowledge in workshops and seminars.

Organization: Paixon GmbH

Contact: christoph.sueess@paixon.ch

Presentation

Christof Traber, Schamai Safra, Bleicke Holm, Dominic Schnyder, and Philipp Lichtenberg

Text-to-Speech (TTS) for Seven Swiss German Dialects

Abstract:

SlowSoft GmbH develops text-to-speech (TTS) components for Swiss German dialects. By mid-2019, TTS for seven major dialects (GR, ZH, BE, LU, SG, BS, VS) will be available, each at first with a single voice (three male, four female voices).

Typical applications of Swiss German TTS include speaking aids (e.g., for ALS patients), language learning tools, chatbots, robots, and digital assistants. Even for applications that finally use recorded speech as output (e.g., IVR systems), TTS may provide an excellent means to create prototypes during the development phase of the application.

Whereas the voice generation component (the "back end") of Swiss German TTS is comparable to that of other languages, the NLP component (the "front end") is extremely challenging due to the fact that there is no standardized way of writing Swiss German. A desirable choice of textual input for most TTS applications would therefore be to use Standard German as input, which, however, would require a full-fledged translation from Standard to Swiss German. A somewhat easier solution is the use of normalized Swiss German, i.e., a word-by-word representation of the Swiss German content by means of Standard German words. The currently adopted solution lies somewhere between these extremes.

The presentation describes the state of development of the Swiss German TTS system, with a focus on the challenges associated with the input text processing.

Biography: Christof Traber studied informatics and holds a PhD in Electrical Engineering from ETH Zurich. He was/is co-founder of the speech-technology companies SVOX AG (acquired by Nuance Communications in 2011) and SlowSoft GmbH (concentrating on minority languages). His experience and interests are language-independent TTS system architecture and development, currently with a focus on languages without standardized text representation (Swiss German).

Organization: SlowSoft GmbH

Contact: christof.traber@slowsoft.ch, schamai.safra@slowsoft.ch, bleicke.holm@slowsoft.ch, dominic.schnyder@slowsoft.ch, phil.lichtenberg@slowsoft.ch

Presentation

Nils Schaetti

Author Verification in Stream of Text with Echo State Network-based Recurrent Neural Models

Abstract:

This paper evaluates a type of recurrent neural networks (RNN) named Echo State Network (ESN) on a NLP task referred as author verification. In this case, the model has to identify whether or not a given author has written a specific text. We evaluate these models on a difficult task where the goal is to detect the author in a noisy text stream being the result of a collaborative work of an unknown number of authors. We construct a new dataset (denoted SFGram) composed of science-fiction books, novels and magazines. From this dataset we select three authors, published between the 1953 and the 1974, and we evaluate the effectiveness of ESNs with word and character-based representations to detect these authors in a set of 91 science-fiction magazines (containing around 8M of words).

Biography: Nils Schaetti is a Ph.D. student in the field of Machine Learning and Natural Language Processing at the University of Neuchâtel. He received his Bachelor and Master degrees in Computer Science from the University of Franche-Comté (France) in addition to a CFC and an IT technician diploma from the CFPT in Geneva. From 2009 to 2016, he held full-time positions in various private companies and organizations such as the EPFL and the Kudelski Group as a computer engineer and HPC specialist. His research interests are in Machine Learning and Artificial Intelligence, Natural Language Processing and Authorship Attribution.

Organization: Université de Neuchâtel

Contact: nils.schaetti@unine.ch

Presentation

Laura Gander, Antonella Bolt, and Ursula Stäuble

Customer-Driven Architecture for a Chatbot

Abstract:

1.26 million of passengers are travelling in Switzerland with the SBB every day. The customer service before, during and after the journey is therefore more than important. Our customers need a quick and handy way to order tickets, reserve seats and obtain information.

The volume in our customer service center increases year by year. Moreover, a big share of the requests are similar to each other. Accordingly, to keep the volume on a level, reduce cost and let our employees focus on more complex requests, we thought about a chatbot which is a highly qualified and advanced product assistant to help our customers 24/7.

We started a first pilot phase with a basic question-answer bot. The goal was to explore whether our customers would use a chatbot at all and to learn about the kind of questions they would ask. In order to gather more information about our customers we conducted qualitative interviews to get insights in what customers know about using chatbots and what expectations they have.

With the gathered insights of the question-answer bot we created a totally new structured different bot. We revised both architecture and design of the bot and included machine learning modules for more "intelligence". To a certain extent, the bot is now able to capture context and sentiments and recognize when it is time to hand over to an agent in a live chat.

The iterative approach allowed us to integrate customer continuous feedback and leveraging the bot to react more appropriate to the customer's requests. While still being in the process of further improvements, we would like to take this opportunity to present you our findings also in comparison with an almost completely dialog guided chat bot (rule-based). What did work so far, and what did not?

- how to gather insights in user behaviour/interaction with a chatbot
- which use cases are suitable for a chatbot
- why it is so important to give the bot emotions and how to detect them

Biography: Laura Gander is a project manager at SBB where she is responsible for innovation and digitalization in the customer service management. She has a management background with a BSc in Business Administration.

Antonella Bolt is product and innovation manager at SBB contact center and is responsible for various projects aimed at automating the increasing number of customer requests. She has a management background with a BSc in Business Administration and a wide background in customer service.

Ursula Stäuble is working in the SBB IT-Department as a Senior IT Project Manager, Business Analyst and Teamleader. She has taken the responsibility for several projects for the Contact Center and the Customer Service.

Organization: SBB AG

Contact: laura.gander@sbb.ch, antonella.bolt@sbb.ch, ursula.staebule@sbb.ch

Presentation

Stephan Wick and Samuel Läubli

Neural Machine Translation for Increased Human Translation Efficiency at Migros Bank

Abstract:

Migros Bank translates more than one million words per year from German to French and Italian. To support their internal team of human translators, the bank decided to develop an in-house neural machine translation system as the use of online services was ruled out for quality and security reasons. This system has been operational on Migros Bank's own infrastructure since mid-2018, meaning that any new email, website or marketing text can be translated automatically without ever leaving the company's network.

In this presentation, we summarise lessons learnt from both a business and technical perspective. We also present results from a blind comparison of the Migros Bank system with DeepL, and results from a productivity study with the internal translators. The latter shows that human translators are 35% faster when using Migros Bank's machine translation system.

Biography: Stephan Wick is COO at Migros Bank. He holds an Msc in Computer Science and Economy from the University of Zurich. Before joining Migros Bank in 2004 he worked as Software Engineer and Project Manager at SIX, USB and Credit Suisse. Samuel Läubli is Partner and CTO at TextShuttle. He holds an MSc in Artificial Intelligence from the University of Edinburgh, and pursues a PhD in Machine Translation at the University of Zurich.

Organization: Migros Bank AG and TextShuttle AG

Contact: Stephan.Wick@migrosbank.ch, laeubli@textshuttle.ai

Presentation

Deepanwita Datta, Manajit Chakraborty, and Aavek Biswas

Your Click Matters: Enhancing Click-based Image Retrieval Performance through Collaborative Filtering

Abstract:

Image retrieval has been an active research area since the early days of computing. While ensemble, multimodal and hybrid methods coupled with learning has seen an upward surge replacing unimodal, heuristic-based methods; a rather new offshoot has been to identify new features associated with images on the web. One such feature is the 'click count' based on the clicks an image or its corresponding text gets in response to a query. Previous state-of-the-art methods have tried to exploit this feature by using its raw count and machine learning. In this paper, we build on this idea and propose a new collaborative filtering based technique to employ the click-log of users from the web to better identify and associate images in response to either a text or an image query. Experiments performed on a large scale publicly available standard dataset having genuine click logs from actual users corroborate the efficacy and significant increase in efficiency of our approach.

Biography: Deepanwita Datta completed her Ph.D. at the Department of Computer Science and Engineering at Indian Institute of Technology (BHU), Varanasi in 2018. Her thesis was on Multimodal Information Retrieval.

Manajit Chakraborty is a PhD student at the Faculty of Informatics, University of Lugano. Manajit does research in Information Retrieval, Text Mining and Social Network Analysis. Aavek Biswas is a Master's student in Computer Science at the University of California, San Diego. His research interests include Machine Learning and its applications in Computer Vision.

Organization: Norwegian University of Science and Technology, Università della Svizzera italiana, and University of California San Diego

Contact: ddatta.rs.cse13@itbhu.ac.in, chakrm@usi.ch, a4biswas@ucsd.edu

Presentation

Fabio Rinaldi, Philipp Kuntschik, Gottowik Jürgen, Mathias Leddin, Raul Rodriguez Esteban, Albert Weichselbraun, Tilia Ellendorff, Nico Colic, and Lenz Furrer

MedMon: Social Media Analytics for a Healthcare Application

Abstract:

We present MedMon, a collaborative project supported by InnoSuisse, focused on the analysis of social media to capture the exchange of biomedical information among providers, patients, scientists, pharmaceutical companies and other healthcare stakeholders.

The project aims at developing a text analytics platform specifically tailored to social media and to be used within a major pharmaceutical company. Two application scenarios under consideration are described below:

1. Clinical trials only capture parameters predefined in protocols. Through social media it is possible to gather patient discussions on disease burden, co-morbidities, quality of life and activities of daily living to identify aspects not covered by trial protocols to reduce potential study burden for patients.
2. Emerging information relevant to different disease areas needs to be monitored, filtered and aggregated, however patient blogs, social media and news channels cannot be queried systematically with existing tools. MedMon aims at providing more efficient means to query these data sources. Using social media analysis to gain improved patient insights has the potential to save significant amounts of resources in the process of drug development, and improve patient outcomes measures.

The project partners have developed a fully functional prototype of the analytical platform and are researching technical advancements to refine the accuracy of the delivered results.

Biography: Fabio Rinaldi is a senior researcher and lecturer at IDSIA, Lugano, and at the Institute of Computational Linguistics, University of Zurich. He leads a research group focusing on biomedical text mining (<http://www.ontogene.org/>). Currently he is principal investigator in an SNF project and in an InnoSuisse project, and partner in several other projects. He has more than one hundred publications, and long experience in the presentation of research results.

Organization: University of Zurich, HTW Chur, and Roche

Contact: fabio@ontogene.org, Philipp.Kuntschik@htwchur.ch, furrer@cl.uzh.ch, juergen.gottowik@roche.com, mathias.leddin@roche.com, ellendorff@ifi.uzh.ch, colic@ifi.uzh.ch, raul.rodriguez-esteban@roche.com, Albert.Weichselbraun@htwchur.ch

Presentation

Aaron Richiger

Automated Data Extraction from Documents using Machine Learning

Abstract:

About 80 percent of all business-relevant data is unstructured. Given the variety in layouts and changing contexts of information in a document, data fields or texts are still manually copied for further processing. This is time-consuming, error-prone and costly. Bringing semi-structured and unstructured data into a machine-readable format is therefore an ongoing challenge and prerequisite also in the field of text analytics.

Over the last years, we developed a document extraction platform – MINT.extract – that allows us to access relevant data (text, images, etc.) from any type of document in a highly efficient and flexible way. In this talk, we present a customer case to process purchase orders using machine learning. Besides automated extraction of information like order number, article number, ordering date, and the number of items ordered, we also perform validation steps. The system runs as a service 24/7 and is accessible from all our client's global locations. As a further benefit, the learning system is highly scalable given the small set of training data required to get the accuracy scores above 95%. We will share some technical insights on how we achieved this. Finally, an intuitive user interface allows users to train their own learning system independently and the architecture enables a seamless integration with existing ERP systems. Naturally, this approach can be applied to insurance policies, invoices, official documents or entire book series as well as other document types.

Biography: Aaron Richiger is a passionate entrepreneur and gifted software engineer working full-time at turicode AG, which he co-founded. He employs his persistent scientific curiosity to develop novel software solutions and is responsible for Machine Learning within turicode. He completed both his bachelor's and master's degree in computer science at ETH Zurich.

Organization: turicode AG

Contact: aaron.richiger@turicode.com

Presentation

Gion Linder

Success of ASR (Automatic Speech Recognition) in the Field of Subtitling

Abstract:

Subtitling is a labour-intensive activity and thus expensive. Therefore, there is a strong pressure to replace this activity with automatic speech recognition and post-editing. But does it really work? Do the automatic subtitles meet the quality requirements? And how should the correction loop be designed so that the texts created with automatic speech recognition can be corrected efficiently? The subsequent proofreading process is almost as important as the quality of automatic speech recognition, since manpower for proofreading is the biggest cost factor.

Biography: Gion Linder is currently Head of Access Services at SWISS TXT. After studies in marketing Gion Linder worked as a consultant for different new media projects in the eighties and nineties. He worked as a content-manager for a publisher, as a product-manager for a cable-company and as a project-manager for a mobile carrier. In 2007 he joined SWISS TXT, a company of the national public broadcaster, where he strongly increased the number of services for television and third-party entities in all national languages. Meanwhile he is a member of the management board. Gion Linder is also chairman of the Eurovision Access Services Experts Group.

Organization: SWISS TXT AG

Contact: Gion.Linder@swisstxt.ch

Presentation

Gabriel Luthier and Andrei Popescu-Belis

PLACAT: A User-friendly Question Answering System for Smart Speaker Devices

Abstract:

The goal of the PLACAT project is to combine the conversational capacities of chatbots, which can be trained on large amounts of non-directed dialogue, with the question answering (QA) capability of dedicated systems. The result is demonstrated as a virtual assistant on a commercial smart speaker, using its automatic speech recognition capabilities. The virtual assistant includes three key technical components, two from the state of the art and one innovative.

The key innovation is a dialogue controller that directs input utterances, depending on their recognized dialogue act, either to a robust chatbot (if the utterance is labeled as interaction-related, such as greetings or chit-chat) or to a powerful question answering system (if the utterance is recognized as a request for information). The controller is a classifier trained on a database composed of questions from chat corpora and from QA ones. The chatbot uses recent models based on reinforcement learning trained on chat corpora, while the QA system uses the recent BERT representations of words in context (from a Transformer neural network) trained on a large QA database (SQuAD).

The demonstrator uses the Google Home speaker. It forwards spoken actions to the controller on the DialogFlow platform, which then connects to our servers for each of the dialogue components. The PLACAT platform is intended as a showcase for the development of commercial virtual assistants endowed with a strong capacity for question answering, particularly over corporate documents such as technical manuals, but at the same time able to sustain user-friendly conversations.

Biography: Gabriel Luthier is a research assistant at HEIG-VD / HES-SO and the main developer for the PLACAT project. In his BSc thesis, he designed a dialogue chatbot for a major tourism office. Andrei Popescu-Belis is a professor at HEIG-VD / HES-SO, a lecturer at EPFL, and a researcher in human language technology, specializing in machine translation and information retrieval. He has published over 150 refereed papers in the field.

Organization: HEIG-VD / HES-SO

Contact: gabriel.luthier@heig-vd.ch, andrei.popescu-belis@heig-vd.ch

Presentation

Andrew Marritt

Building an Inductive Coding Service for Employee Feedback

Abstract:

Large firms are increasingly using feedback as a way of collecting information on ideas on a wide-range of business and employee topics. Workometry, a 2018 Gartner Cool Vendor, is a service for large employees that enables them to quickly make sense of large volumes of multi-language employee feedback.

Feedback, often provided as sentence-fragments, needs to be understood in the context of the question. Decision makers have a strong preference for identified themes to themselves be reasonable answers to the question rather than noun-based topics. We therefore build inductive, question and organization-specific models aiming do this within a few hours. Our approach uses a human-in-the-loop to improve accuracy to perform a service with similar performance to human coders but at a fraction of the resource requirements.

In this presentation I discuss how we built and continually refine the service. I will discuss some of the challenges we overcame and how small changes in the survey employees are presented can improve (or decrease) the ease of building successful models.

Biography: Andrew Marritt is the CEO of OrganizationView, a St. Moritz based pioneer in applying data to workforce challenges. He has held senior global HR roles in firms including Alstom, UBS and Reuters following an early career in management consulting. Andrew has a BSc. in Economics. He learnt to program in the late '70s on a computer built from a kit. He teaches HR Analytics at FHNW, was the Chair of the largest HR analytics conference for their first 3 years and is a member of the CIPD advisory group on HR Analytics.

Organization: OrganizationView GmbH

Contact: andrew.marritt@organizationview.com

Posters

Bitcoin Value and Sentiment Expressed in Tweets

Authors: Bernhard Preisler, Margot Mieskes, and Christoph Becker

Organization: University of Applied Sciences Darmstadt

Contact: preisler.bernhard@gmail.com, margot.mieskes@h-da.de, christoph.becker@h-da.de

The Reader's Feeling and Text-based Emotions: the Relationship between Subjective Self-reports, Lexical Ratings, and Sentiment Analysis

Authors: Egon Werlen, Fernando Benites, Christof Imhof, and Per Bergamin

Organizations: Swiss Distance University of Applied Sciences - FFHS and Zurich University of Applied Sciences - ZHAW

Contact: egon.werlen@ffhs.ch, benf@zhaw.ch, christof.imhof@ffhs.ch, per.bergamin@ffhs.ch

Talos, the Deep Learning Solution of Credit Suisse

Authors: Angeliki Davourli

Organization: Credit Suisse

Contact: angeliki.davourli@credit-suisse.com

The PsyMine Project: Outcomes and Insights

Authors: Tilia Ellendorff, Simon Foster, and Fabio Rinaldi

Organizations: University of Zurich and ISGF

Contact: ellendorff@cl.uzh.ch, simon.foster@isgf.uzh.ch, fabio.rinaldi@uzh.ch

Summarization of German Texts

Author: Dominik Frefel

Organization: University of Applied Sciences Northwestern Switzerland (FHNW)

Contact: dominik.frefel@fhnw.ch

Posters

DeepDTI: Identification of Drug-Target Interactions using Deep Learning Techniques

Authors: Hatem Ghorbel, Fabrizio Albertetti, and Jérôme Moret

Organizations: Haute École Arc Ingénierie - University of applied Sciences of Western Switzerland (HES-SO)

Contact: hatem.ghorbel@he-arc.ch, fabrizio.albertetti@he-arc.ch, jerome.moret@he-arc.ch

Multi-language Multi-label Email Classification at SBB's Customer Support

Authors: Nadja Keidel and Marlyse Lichtsteiner

Organizations: Zühlke Engineering AG and SBB

Contact: nadja.keidel@zuehlke.com, marlyse.lichtsteiner@sbb.ch

How Patient Insights from Social Media Improve a Conceptual Disease Model

Authors: Mathias Leddin, Raul Rodriguez-Esteban, and Jürgen Gottowik

Organization: F. Hoffmann-La Roche Ltd.

Contact: mathias.leddin@roche.com, raul.rodriguez-esteban@roche.com, juergen.gottowik@roche.com

Sentiment Analysis for a Swiss Gig Platform Company

Authors: Ela Pustulka-Hunt and Thomas Hanne

Organizations: University of Applied Sciences Northwestern Switzerland (FHNW Olten)

Contact: elzbieta.pustulka@fhnw.ch, thomas.hanne@fhnw.ch

QUAND - Quantitative Need & Demand Modeling from Reviews

Authors: Ritavan, Sameh Metias, and Harald Koppen

Organization: Steering Lab, Horváth & Partners

Contact: ritavan.saice@gmail.com, samehmetias@gmail.com, HKoppen@horvath-partners.com

Posters

Quality Assessment of Automatic Speech Recognition Systems

Authors: Fabian Germann, Malgorzata Anna Ulasik, and Mark Cieliebak

Organization: ZHAW

Contact: germafa2@students.zhaw.ch, ulasimal@students.zhaw.ch, ciel@zhaw.ch

Matching Human Competences with Semantic Search

Authors: Kurt Wehrli, Massimo Lusetti, and Tanja Samardzic

Organizations: Symsala and University of Zurich

Contact: kurt.wehrli@symsala.com, massimo.lusetti@uzh.ch, tanja.samardzic@uzh.ch

Semantic Systems and Visual Tools for Analyzing Domain-specific Text Corpora and Communication Channels

Authors: Albert Weichselbraun and Philipp Kuntschik

Organizations: University of Applied Sciences, Chur, and Swiss Institute for Information Research, University of Applied Sciences Chur

Contact: albert@weichselbraun.net, Philipp.Kuntschik@htwchur.ch

Email Anonymization in the Insurance Industry

Authors: Anastasios Zouzias, Theus Hossmann, Benjamin Theunissen, and Christian Müller

Organizations: Sqooba (Schweiz) AG and Helvetia Versicherungen Schweiz

Contact: anastasios@sqooba.io, theus@sqooba.io, benjamin.theunissen@helvetia.ch, christian.mueller@helvetia.ch

News from dialektkarten.ch

Author: Yves Scherrer

Organization: University of Helsinki

Contact: yves.scherrer@helsinki.fi

Demos

Demonstration of LARA, a Tool to Help Develop Foreign Language Reading Skills

Authors: Elham Aklaghi, Branislav Bedi, Cathy Chua, Hanieh Habibi, and Manny Rayner

Organizations: Ferdowsi University of Mashhad, University of Iceland, Ex-PB, and University of Geneva

Contact: elhamakhlaghi80@gmail.com, brb19@hi.is, cathy@pioneerbooks.com.au, Hanieh.Habibi@unige.ch, Emmanuel.Rayner@unige.ch

Using AI Speech and Text Analytics to Uncover Valuable Insights for Contact Centre Compliance and Complaints

Authors: Matteo Amore, Federica Chierici, and Mariella Borghi

Organization: CELI Language Technology

Contact: matteo.amore@celi.it, federica.chierici@celi.it, mariella.borghi@celi.it

Corporate Sustainability Assessment with Modern NLP

Authors: Ursin Brunner, Arthur Habicht, Dorian Tanase, and Pascal Wyss

Organization: ti&m AG

Contact: ursin.brunner@ti8m.ch, arthur.habicht@ti8m.ch, dorian.tanese@ti8m.ch, pascal.wyss@ti8m.ch

Adobe Sense - Customer Connectivity Redefined

Authors: Rupak Chakraborty and Kapil Arora

Organization: Adobe Inc

Contact: rupak97.4@gmail.com, karora@adobe.com

Automatic Transcription of Interviews

Authors: Mark Cieliebak, Nicolas Eckhart, Mathias Marxer, and Malgorzata Ulasik

Organization: Zürcher Hochschule für Angewandte Wissenschaften

Contact: ciel@zhaw.ch, eckhnic@students.zhaw.ch, marxemat@students.zhaw.ch, ulasimal@students.zhaw.ch

Demos

Historical Media Monitoring with impresso

Authors: Maud Ehrmann, Matteo Romanello, Frédéric Kaplan, Marten Düring, Estelle Bunout, Paul Schroeder, Daniele Guido, Thijs van Beek, Simon Clematide, Gerold Schneider, Phillip Benjamin Ströbel, Martin Volk, and Andreas Fickers

Organizations: Ecole Polytechnique Fédérale de Lausanne, C2DH, and University of Zurich

Contact: maud.ehrmann@epfl.ch, matteo.romanello@epfl.ch, frederic.kaplan@epfl.ch, marten.during@uni.lu, estelle.bunout@uni.lu, paul.schroeder@uni.lu, daniele.guido@uni.lu, thijs.vanbeek@uni.lu, siclemat@ifi.uzh.ch, gschneid@ifi.uzh.ch, pstroebel@cl.uzh.ch, volk@cl.uzh.ch, andreas.fickers@uni.lu

Deep Learning of Charisma

Authors: Philip Garner, Olivier Bernet, Dimitra Loupi, Dominic Rohner, and John Antonakis

Organizations: Idiap Research Institute and University of Lausanne

Contact: pgarner@idiap.ch, Olivier.Bernet@idiap.ch, loupidimitra@gmail.com, dominic.rohner@unil.ch, John.Antonakis@unil.ch

Transfer Learning for Feedback Analysis: Demo of codit.co

Authors: Maurice Gonzenbach and Pascal de Buren

Organization: Caplena

Contact: maurice@caplena.com, pascal@caplena.com

Predict Mortality Rates from Twitter

Author: Ralf Grubenmann

Organization: SpinningBytes AG

Contact: rg@spinningbytes.com

Simplify AI for Business Users

Author: Dr. Hans-Peter Göllich

Organization: Dydon AG

Contact: hpguellich@dydon.net

Demos

A Machine Learning-based Solution for Contract Analytics

Author: Franz Kögl

Organization: IntraFind Software AG

Contact: franz.koegl@intrafind.de

MarketMiner: A Platform for International Web-based Market Intelligence Facilitating Strategic Business Decisions

Author: Janna Lipenkova

Organization: Anacode

Contact: janna.lipenkova@anacode.de

A Framework for Question-Driven Text Corpus Exploration

Authors: Linus Metzler, Nadina Siddiqui, Dominik Steiner, Gezim Zeneli, Nathan Lepori, Jan Albert, Fernando Benites, Don Tuggener, Pius von Däniken, and Mark Cieliebak

Organizations: Zurich University of Applied Sciences (ZHAW)

Contact: linus.metzler@gmail.com, nadinasiddiqui@msn.com, steindo1@students.zhaw.ch, zenelgez@students.zhaw.ch, lepornat@students.zhaw.ch, alberja1@students.zhaw.ch, benf@zhaw.ch, tuge@zhaw.ch, vode@zhaw.ch, ciel@zhaw.ch

Biomedical Text Comprehension using Transfer Learning and Semantic Coherence

Authors: Marilena Oita, Fatma Oezdemir-Zaech, and Fabien Pernot

Organization: Novartis

Contact: marilena.oita@novartis.com, fatma.oezdemir-zaech@novartis.com, fabien.pernot@novartis.com

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

Authors: Jürgen Schwärzler and Philipp Thomann

Organization: D ONE

Contact: juergen.schwaerzler@d-one.ai, philipp.thomann@d-one.ai

Demos

Text Analytics for Keyword Research in Online Marketing – Challenges & Solutions

Author: Christine Spies

Organization: Dept Digital Marketing AG

Contact: christinespies@web.de

Automatic Didactic Feedback for Student Text through Natural Language Understanding

Authors: Anette Hunziker and Samuel Portmann

Organization: Taskbase

Contact: anette@taskbase.com, samuel@taskbase.com

Show Me Your Documents and I'll show You Your Knowledge Graph - Semantic Search in "Unknown" Domains

Authors: Alexandros Paramythis and Doris Paramythis

Organization: Contextity AG

Contact: paramythis@gmail.com, doris.paramythis@contextity.ch



We turn data into valuable insights
with **Natural Language Understanding**

SpinningBytes develops software solutions for Automatic Text Understanding, based on Artificial Intelligence, Machine Learning algorithms, Natural Language Generation and Text Analytics. We build the bridge between research and industry.

Our Services:

- **Onboarding Workshops**
- Rapid Prototypes
- **Ready-to-use Software Solutions**
- Data Analysis and Statistics
- **Training and Education**



SpinningBytes AG

Albanistrasse 20 · 8400 Winterthur · +41 (0)44 500 25 02

www.spinningbytes.com · info@spinningbytes.com

Sponsors



Who we are.

DTI AG based in Wil, Switzerland, has been an independent system integrator and information management expert for more than 25 years. The company focuses on the preparation & analysis of large amounts of data and the provision of context-related information.

We offer tailor-made & highly flexible solutions and focus on the cutting-edge technologies such as: Natural Language Processing (NLP), Machine Learning, Artificial Intelligence (AI), Linguistic, Universal Search and much more.

What we do.

We support. We design. We integrate. We shape comprehensive solutions from different software components of our technology partners for our customers. As individual as their demands.

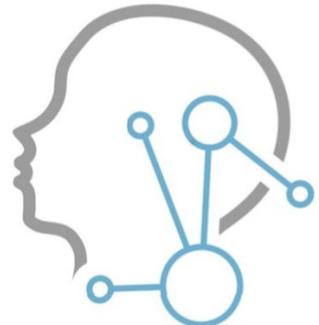
DTI AG | Flawilerstrasse 27 | CH-9500 Wil/SG | T+41 71 929 77 77 | info@dti.ch | www.dti.ch

eonum

We are a Swiss IT company, specializing in Natural Language Processing and Deep Learning. Our main focus is providing solutions to the healthcare sector, specifically in the field of medical coding and documentation. Since the introduction of the DRG system in Switzerland, we have been collecting data and developing tools to provide our customers with forecasts and analytics.

Our Services

- Machine Learning projects with medical text and medical coding
- Data and Market Analysis of the Swiss healthcare sector
- Forecasting medical outcomes
Data-driven
- Data-driven software solutions
- Workshops and training sessions



Visit us on our website www.eonum.ch



die Mobiliar

Die Mobiliar is Switzerland's oldest private insurer, and the market leader in several major insurance segments. Every third household and every fourth company in Switzerland is insured with us.

Our presence in over 160 locations throughout the country cements our reputation as Switzerland's most personal insurance company.

The Data Science & Analytics team at Die Mobiliar works on solving problems centered around insurance data. We combine our expertise on machine learning, text mining, natural language processing, image analysis, and recommender systems with Die Mobiliar's nearly 200-year-old expertise on the insurance business.

The result: analytics solutions that optimise processes and support decision-making for Die Mobiliar and the wider insurance industry.



Contact:
Dr. Mirco Rossi
Head of Cognitive Computing &
Disruptive Analytics
mirco.rossi@mobi.ch
www.mobiliar.ch

Sponsors



Thanks to Swiss technology and artificial intelligence, recapp IT's unique speech recognition works for Swiss languages and dialects.

www.recapp.ch

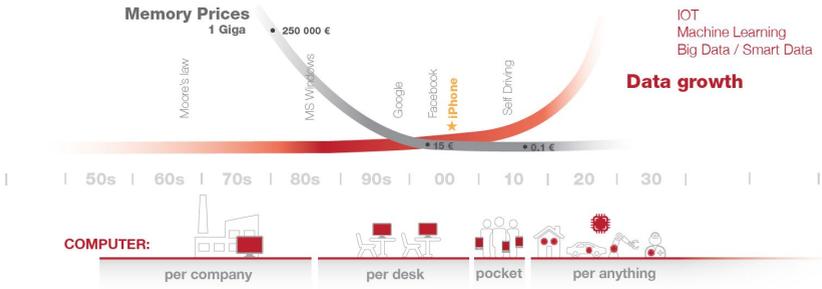
contact@recapp.ch

Artificial intelligence.
And human excellence.



Automated texts are becoming the new normal. But who wants to be normal? Truly great copy isn't made of data points. It's made of sweat, joy and damn good ideas. Let's make it happen. Together. **supertext.ch**

DATA DRIVES DIGITIZATION



DATA DRIVEN VALUE CREATION

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

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



SwissText 2020 is already in the making!

Next year's edition will be collocated with Konvens. Want to stay in the loop? Sign up for the SwissText newsletter at <http://www.swisstext.org/maillinglist!>

Partner

Swiss Association for Analytics



The Swiss Association for Analytics (SAA) has been created in 2012. It is the very first Swiss group entirely dedicated to predictive and descriptive analytics. Our main objective is to raise awareness of Swiss companies to the benefits of analytics. By analytics, we mean the

use of data mining and machine learning algorithms for data-driven decision making.

Whether the domain of your company is banking, finance, pharma, e-commerce or telecom doesn't matter. To benefit from analytics, companies need to have data, tools and know-how. The SAA can provide support in these three areas.



The SAA is a non-lucrative organization with the following objectives:

- Promote analytics within Switzerland
- Show the added value of analytics to Swiss companies
- Provide networking facilities for practitioners
- Exchange with other associations having related objectives

We have several means to achieve the above-mentioned objectives. We manage a LinkedIn group (www.linkedin.com/groups/4586163) in which we discuss topics such as trends, challenges, case studies, events and job offers. We also organize 3-4 analytics events each year (www.meetup.com/swiss-analytics). Our events regularly gather around 100 people in analytics. We also publish a printed magazine dedicated to Analytics, twice a year (www.swiss-analytics.com/magazine). If you are interested to be a sponsor, speaker or author, feel free to contact us at info@swiss-analytics.com.



School of Computer and Communication Sciences
Machine Learning and Optimization Laboratory

Research:

Our main research directions:

- Machine Learning
- Optimization
- Text Understanding
- Unsupervised and Weakly Supervised Learning

We study unsupervised learning of general-purpose representations (features) for text (words and documents) such as [sent2vec](#), and other data modalities including time series (e.g. [CNN based models](#)).

As a second direction, we develop and analyze new optimization algorithms for training machine learning models. Last but not least, our interest is in scalable distributed and decentralized training algorithms and systems for classic ML models as well as deep learning, and open-source implementations and [benchmarking](#) of such systems.

Further Information:

Detailed information about our research, open positions and course materials can be found here: <https://mlo.epfl.ch>

Partner



HAUTE ÉCOLE
D'INGÉNIERIE ET DE GESTION
DU CANTON DE VAUD

www.heig-vd.ch

Hes·SO

Haute Ecole Spécialisée
de Suisse occidentale
Fachhochschule Westschweiz
University of Applied Sciences and Arts
Western Switzerland

Institute for Information and Communication Technologies (IICT)

The School of Management and Engineering Vaud (in French, Haute École d'Ingénierie et de Gestion du Canton de Vaud, HEIG-VD) is one of several schools within the University of Applied Sciences and Arts of Western Switzerland (HES-SO). As all the eight Swiss universities of applied sciences, the main missions of HEIG-VD include basic and continuing education, applied research and development, and services to other institutions. HEIG-VD has more than 2000 students enrolled in ten Bachelor programs, and also contributes to three HES-SO Masters programs. HEIG-VD is located on three sites in the town of Yverdon-les-Bains, one of them being Y-Parc, the first and largest technology park in Switzerland.

The Institute for Information and Communication Technologies (IICT) is the largest R&D institute of HEIG-VD. Its activities pertain to software engineering, telecommunications and IT, with particular emphasis on intelligent data analysis, security, and advanced communication systems. The IICT carries out about 50 research projects every year, often in collaboration with industrial partners. Over the past decade, the IICT has created four startup companies. The IICT has currently about 60 members, including 18 professors and several PhD students in co-supervision with other universities. One of its most promising research directions is machine learning and big data, with applications to sensor data analysis, biomedical engineering, human computer interaction, information retrieval, natural language processing, and machine translation.

HEIG-VD | Route de Cheseaux 1 | CP 521 | 1401 Yverdon-les-Bains

+41 (0) 24 557 63 30 | <https://heig-vd.ch> | <http://iict.heig-vd.ch/>

Partner

Swiss Distance University of Applied Sciences



Created in 1998, the Swiss Distance University of Applied Sciences (Swiss Distance UAS) offers nationally recognised Bachelor and Master's degrees as well as a range of further education courses in economics, information technology, law and health.

In 2004, the Swiss Distance UAS became an affiliated school of the Scuola Universitaria Professionale della Svizzera Italiana (SUPSI). The Swiss Distance UAS offers degree courses in applied science based on blended learning. This means that about 80% of the training takes place through supervised distance study, while the remaining 20% is done at the regional centres in Zurich, Berne, Basel and Brig. By offering innovative courses that combine different types of e-learning with traditional education methods, and by training professionals without removing them from their jobs, the Swiss Distance UAS has closed a gap in the Swiss educational system. Thanks to this strategy it achieves a high degree of specialisation in each field of study.

In its three research institutes, the Swiss Distance UAS engages in applied research covering web science, management, innovation and e-learning that complies with all national Swiss education standards:

Institute for Research in Open-, Distance- and eLearning (IFeL)

The Institute for Research in Open-, Distance- and eLearning (IFeL) investigates the use of modern technology in learning. The interdisciplinary team of the IFeL consists of social and computer scientists and collaborates with both national and international partners. They apply a design-based research methodology and focus on emotions in reading and learning. Since autumn 2016, the institute has been home to the **UNESCO Chair for Personalised and Adaptive Distance Education**. In this context, it investigates the framework conditions and implementation concepts for personalised and adaptive learning.

The Institute for Management and Innovation (IMI)

The IMI research in the area of management and innovation. The institute cooperates with enterprises, organizations and groups as well as partner institutes from other colleges.

The core competencies of the IMI is in the thematic areas of innovation management, Organizational Intelligence: How to foster interdisciplinary collaboration to increase innovation organizational intelligence, interdisciplinary collaboration, organizational capabilities, and absorptive capacity. Research in these areas has been conducted, published in scientific journals and practitioner magazines, and presented at various conferences. The research have a strong background in qualitative research methods and is experienced in collaborating with management practitioners.

Laboratory for Web Science (LWS)

The Laboratory for Web Science (LWS) is a joint research unit of the FFHS and the Information Systems and Networking Institute (ISIN) of the SUPSI. The LWS focuses on the field of data science and it closely collaborates with industry partners on research projects to facilitate innovation and technology transfer.

ffhs.ch/en

Partner

Lucerne University of
Applied Sciences and Arts

HOCHSCHULE LUZERN

Information Technology

FH Zentralschweiz

Lucerne School of Information Technology

The Lucerne University of Applied Sciences and Arts is the first Swiss university of applied sciences to have its own School of Information Technology. More than 600 students are enrolled in the bachelor's degree programs in Digital Ideation, Computer Science, Information & Cyber Security, International IT Management and Business Information Technology and master's degree programs in Engineering, Business Information Technology and Specialized Media and IT Teaching and Learning Methods at the Zug-Rotkreuz campus. The School of Information Technology is focusing its research on new areas of interest, ranging from artificial intelligence, machine learning and mobile and smart systems all the way to blockchain, smart contracts and visual computing. Numerous partners from the public and private sector benefit from the project skills and professional expertise of the researchers.

Continuing and executive education programs include the Master of Advanced Studies and Certificate of Advanced Studies in four attractive specialist areas. Courses, seminars and conferences lasting one or more days are also available. The Lucerne School of Information Technology also has outstanding links to domestic and foreign universities as well as to partners from all areas of economic, administrative, cultural and social life. Thanks to its range of programs, the School of Information Technology offers added value for business, society and culture.

For more information about this institution, please contact Prof. Dr. René Hüsler, Dean of the Lucerne School of Information Technology (rene.huesler@hslu.ch).

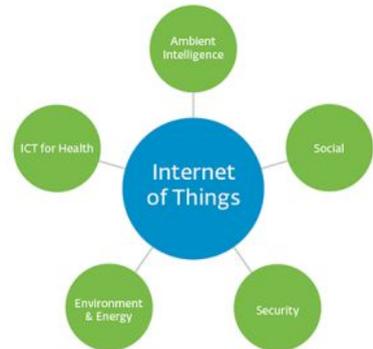
SUPSI

Information Systems and Networking Institute

The Information Systems and Networking Institute (ISIN) is part of the Department of Innovative Technologies at SUPSI. ISIN carries out teaching (Bachelor, Master and Continuing education) and applied research activities in the information and communication technology domain with a special focus of Internet of Things applications.

The robust research portfolio of the Institute spans many disciplines, including:

- Data and semantic analysis
- Multimedia processing
- Natural User Interfaces
- Cyber-security
- Pervasive communication



The wide range of competences is mostly exploited in Internet of Things projects involving local companies and other organizations.

Since it was established, the institute has considerably grown in terms of people, projects and applied research activity reaching about 40 employees and 3 Mio CHF yearly in applied leading-edge research projects. In the last few years dozens of KTI, EU and SNF projects have been acquired and completed.

SUPSI - Dipartimento tecnologie innovative

Istituto sistemi informativi e networking

Galleria 2, Via Cantonale 2c

CH-6928 Manno

T +41 (0)58 666 65 83

F +41 (0)58 666 65 71

isin@supsi.ch



Institute of Computational Linguistics

Text Technologies and Phonetics

Computational Linguistics investigates how human language is used as a means of transmitting, storing and processing information, and how these processes can be modeled on a computer and made available to specific applications. Searching information in the World Wide Web, analyzing texts in blogs and forums to gain insight in people’s opinions, automatic text summarization or machine translation – Computational Linguistics attempts to make information available for our knowledge-based society.

The phonetics group, which joined the Institute in 2018, carries out fundamental research on different aspects of phonetics and speech sciences, in particular in the field of forensic phonetics.

Study

Computational Linguistics combines linguistics and information science and is the right choice for everyone that is interested in both areas.

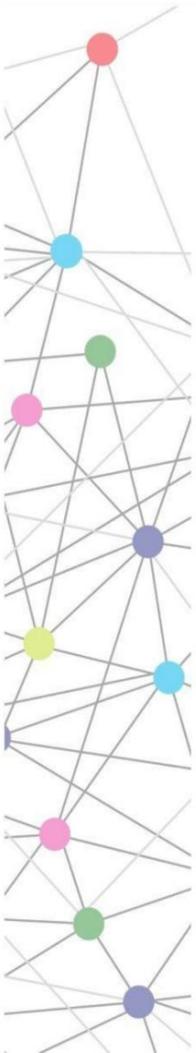
Research

Research topics of our institute include: Multilingual Text Analysis
Machine Translation
Sentiment Analysis and Opinion Mining
Automated Media Analysis

Biomedical Text Mining
Phonetics and speech sciences
Forensic phonetics

Further information

Detailed information about our research and studying Computational Linguistics can be found at: www.cl.uzh.ch



Partner



What to do with 70+ data scientists?

The ZHAW Data Science Laboratory, Datalab, is a virtual research group spanning across the departments of one of Switzerland's biggest universities of applied sciences. It brings together all researchers engaged in the aspects of data science under one roof for collaborative research and industry projects.

Founded in early 2013 as one of Europe's first groups dedicated to data science, the ZHAW Datalab currently comprises more than 70 researchers from as diverse areas as law, analytics, computer science, and entrepreneurship. It is one of the leading data science research centers in Switzerland and beyond.

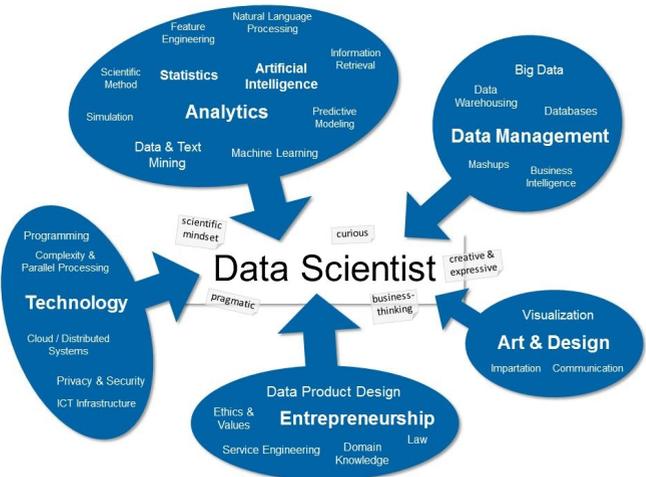
Our offer: R&D projects

The ZHAW Datalab offers the perfect team composition for any individual data science project. From our pool of domain experts, we form interdisciplinary project teams with the right mix of methodical and industry expertise. We are used to work together and cross departmental borders. Possible project settings are third party funded or directly funded R&D projects as well as student's thesis projects, feasibility studies and consulting. Typical project durations vary from a few weeks to several years. We adopt our processes, tools and technologies to our partner's needs.

Our service: education & community building

The ZHAW Datalab designed one of the first dedicated data science curricula in Europe: The *Master of Advanced Studies in Data Science* for professional education. Our associates also commit considerable amounts of time to various Bachelor, Master and Ph.D. degree programs, thus sustaining a continuous idea interchange between projects, students and industry.

We are committed to the Swiss data science community by means of several events and conferences we (co-)organize. One example is the "Swiss Conference on Data Science" series since 2014. We highly value this chance for idea exchange, trend spotting, and passing on of lessons learned.



Get in touch

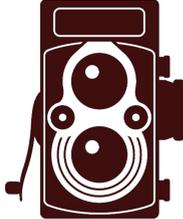
www.zhaw.ch/datalab | datalab@zhaw.ch | [@DataScienceCH](https://twitter.com/DataScienceCH)

Partner



LOHUBER.COM
videographer
photographer

Partner



Cloudia Chen Photography

Cloudia Chen - Destination Photographer around the world, Photographer of Zurich Tourism Office, winner of the public voting 1st prize at Swiss Photo Award 2018.

Cloudia is a portrait photographer specialized in professional headshots and corporate portraits. Before settling in Zurich, she lived in Mainland China, Hong Kong, London and South of France. Her photos and articles have been published in various social media and popular travel magazines.

Website: www.cloudiachen.com

Email: info@cloudiachen.com



Feedback



<https://www.swisstext.org/2019/feedback>

Impressum

Conference Chair

Prof. Dr. Mark Cieliebak

Program Committee

See next page.

Organization Committee

Andrea Aguilar Ibáñez
Dr. Fernando Benites
Jan Deriu
Pamela Hunziker
Manuela Hürlimann
Dr. Don Tuggener
Ada Wan
Cornelia Wichser
Annette Zani

SwissText is organized by Zurich University of Applied Sciences (ZHAW) and co-organized by the Swiss Alliance for Data-Intensive Services. It is supported by the Swiss Innovation Agency (Innosuisse) and industrial sponsors.

Contact

Prof. Dr. Mark Cieliebak
+41 (0) 58 934 72 39
<https://www.swisstext.org>
info@swisstext.org

Program Committee

Swiss Track

- Noah Bubenhofer, Zurich University of Applied Sciences (ZHAW)
- Hatem Ghorbel, HEG – Haute école de gestion Arc
- Martin Jaggi, Ecole Polytechnique Fédérale de Lausanne (EPFL)
- Roberto Mastropietro, Scuola universitaria professionale della Svizzera italiana (SUPSI)
- Thoralf Mildenerger, Zurich University of Applied Sciences (ZHAW)
- Andrei Popescu-Belis, School of Business and Engineering Vaud (HEIG-VD)
- Fabio Rinaldi, University of Zurich (UZH)
- Maria Sokhn, HEG – Haute école de gestion Arc
- Jürgen Spielberger, Zurich University of Applied Sciences (ZHAW)
- Jürgen Vogel, Bern University of Applied Sciences (BFH)
- Manfred Vogel, Fachhochschule Nordwestschweiz (FHNW)
- Tim Vor der Brück, Lucerne University of Applied Sciences and Arts (HSLU)
- Egon Werlen, Swiss Distance University of Applied Sciences (FFHS)

Scientific Track

- Adrien Barbaresi, Berlin-Brandenburgische Akademie der Wissenschaften (BBAW)
- Roberto Basili, University of Roma Tor Vergata
- Ivo Blohm, University of St. Gallen
- Ernst Buchberger, Medical University of Vienna
- Miriam Butt, Universität Konstanz
- Pascual Cantos, Associate Professor, University of Murcia
- Sylviane Cardey, Centre Tesniere, Université de Franche-Comte
- Ernesto William De Luca, Georg-Eckert-Institute for international Textbook Research
- Gaël Dias, University of Caen Normandie
- Nicola Ferro, University of Padova
- Xavier Gómez Guinovart, Universidade de Vigo
- Martin Jaggi, ETH Zurich
- Manfred Klenner, University of Zurich (UZH)
- Valia Kordoni, Humboldt University Berlin
- Udo Kruschwitz, University of Essex
- Ekaterina Lapshinova-Koltunski, Saarland University
- Roberto Mastropietro, SUPSI
- Alexander Mehler, Goethe University Frankfurt
- Margot Mieskes, University of Applied Sciences, Darmstadt
- Marie-Francine Moens, Katholieke Universiteit Leuven
- Guenter Neumann, German Research Center for Artificial Intelligence (DFKI)
- Johann Petrak, The University of Sheffield
- Hannes Pirker, Austrian Centre for Digital Humanities (ACDH) ÖAW
- Andrei Popescu-Belis, HEIG-VD / HES-SO
- Uwe Quasthoff, University Leipzig
- Fabio Rinaldi, University of Zurich (UZH)
- Paolo Rosso, Universitat Politècnica de València
- Tanja Samardzic, University of Zurich (UZH)
- Yves Scherrer, University of Helsinki
- Gerold Schneider, University of Zurich (UZH)
- Rico Sennrich, The University of Edinburgh
- Kurt Stockinger, Zurich University of Applied Sciences
- Manfred Vogel, Fachhochschule Nordwestschweiz (FHNW)
- Jürgen Vogel, Berner Fachhochschule (BFH)
- Tim Vor der Brück, Lucerne University of Applied Sciences and Arts (HSLU)
- Egon Werlen, Swiss Distance University of Applied Sciences (FFHS)
- Torsten Zesch, University of Duisburg-Essen
- Heike Zinsmeister, University of Hamburg



Gold Sponsor

Swiss Alliance for
Data-Intensive Services

Together we move faster.

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, Natural Language Processing and Predictive Maintenance.

To boost innovation several initiatives exist in the Alliance, e.g. the initiative “From ideas to projects” and the start-up grant scheme. 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)





Schweizerische Eidgenossenschaft
Confédération suisse
Confederazione Svizzera
Confederaziun svizra

Swiss Confederation

Federal Department of Economic Affairs,
Education and Research EAER
Commission for Technology and Innovation CTI
Innovation Promotion Agency

D | ONE



SWISS TXT



zhaw

Google



EPFL

eonum



SUPSI

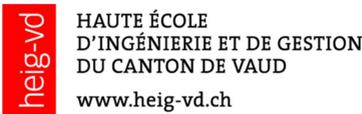


swiss made software



[rec]app

haute école neuchâtel berne jura **arc** ingénierie
www.he-arc.ch



n | w University of Applied Sciences and Arts Northwestern Switzerland



University of Zurich ^{UZH}

die Mobiliar

FFHS

Lucerne University of Applied Sciences and Arts

HOCHSCHULE LUZERN

