



Industrial Applied Data Science @ ZHAW

Zurich University
of Applied Sciences



datalab
www.zhaw.ch/datalab

Prof. Dr. Thilo Stadelmann

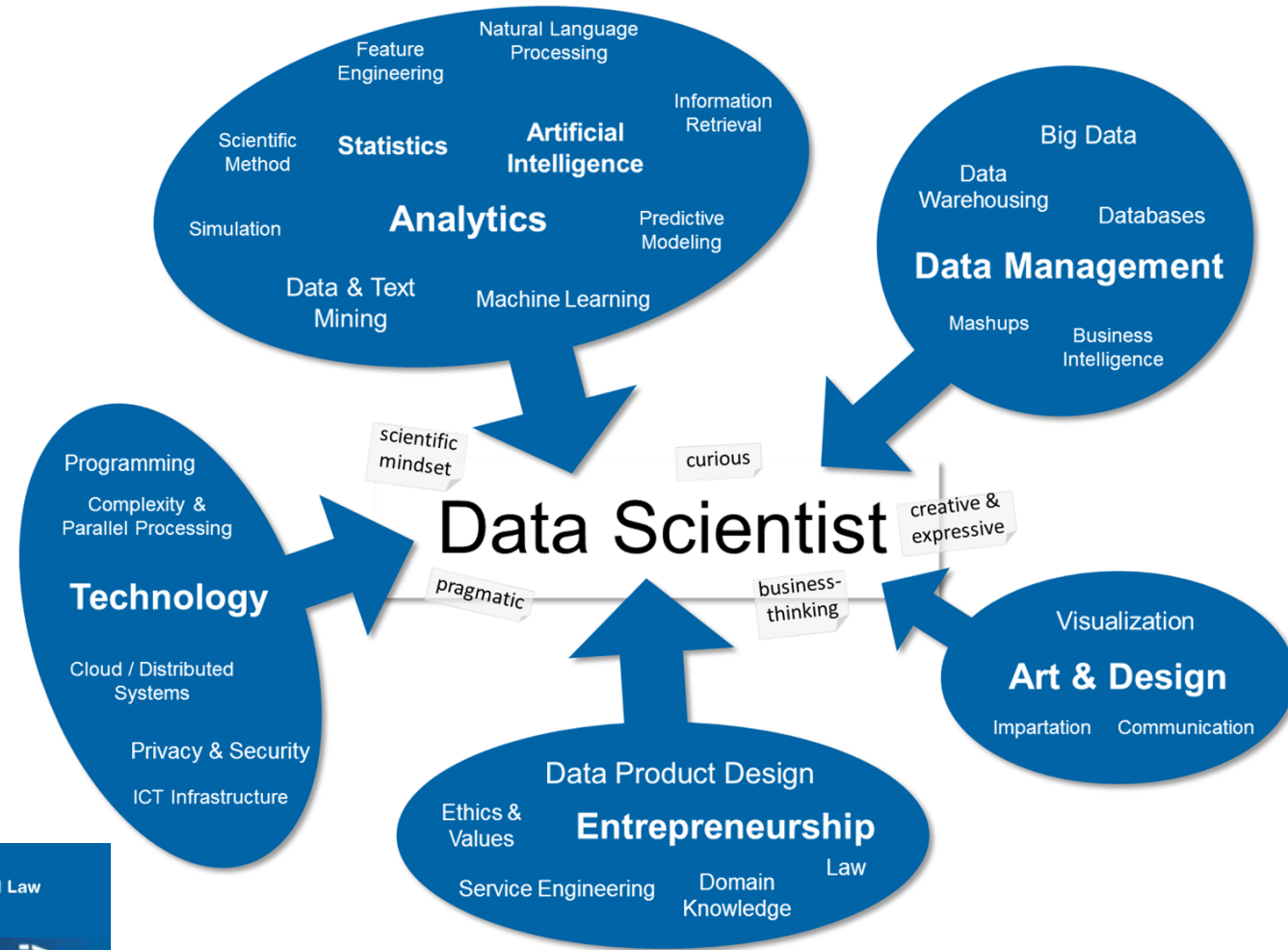
Head of Datalab, the ZHAW Data Science Laboratory

1. Key research topics and people

People: ca. 80 researchers from 3 departments

Digitalisation in practice:

- **Statistical data analysis, predictive modeling**
- **Machine learning**, in particular **deep learning**
- **Information retrieval, text analytics**
- **Big data**, databases & data warehousing
- Information **visualization, visual analytics**
- Business intelligence, decision support
- **Privacy**, security, **data product design**
- **Subject matter expertise:** law, life sciences, ...



zhaw School of Engineering	Life Sciences und Facility Management	School of Management and Law
InIT	IDP	IAMP
	IAS	ZSR



See Stadelmann, Stockinger, Braschler, Cieliebak, Baudinot, Dürr and Ruckstuhl: *Applied Data Science in Europe*. ECSS 2013, Amsterdam.

2. Research projects

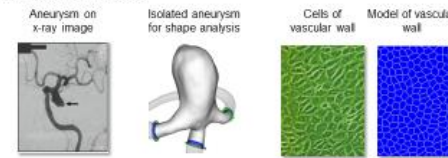
Volume

- > 9 Mio. CHF 3rd party funding in first 4 years

Topics: all of digitalisation

- Industry 4.0 (e.g., CTI project QualitAI)
- E-Health (e.g. SystemsX/SNSF project AneuX)
- FinTech (e.g., CTI project DatFrisMo)
- Mobility (e.g., project Placebook)
- Sustainability (e.g., CTI project EAT-IT CO₂)
- Technology (e.g., SNSF project Bio-SODA)
- Document Analysis (e.g., CTI prj. DeepScore)
- ...

AneuX: Is aneurysm shape significant for its hazardousness?

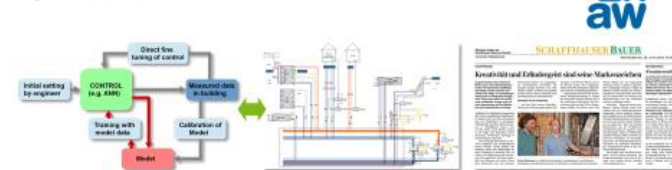


- SystemsX.ch funding: 2M CHF, review by SNSF
- Morphological analysis of aneurysms using machine learning
 - Biologically motivated simulation model of cell wall change
 - Creation of disease model for treatment planning
 - Creation of database of aneurysms
 - Development of tools to analyze clinical (image) data

Partners (co-applicant Sven Hirsch, ZHAW):



Hydrobus: Simulation-based Optimization



- The challenge
- Not enough training data for AI in socio-technological systems

- The project
- Self-adaption of control to time-varying demands in a multi-apartment building using simulations
 - Combined entropy and energy optimization of HVAC-system based on Model Predictive Control
 - Integrates renewable energy technology, social dynamics and scenario-based weather prediction

- The upside
- Enables a Swiss SME to harvest results from modern mathematics, data science and AI
 - Gives science the opportunity to test modern approaches on real-world problems



ARGUS MEDIA BASED INTELLIGENCE

PANOPTES – Automated Article Segmentation of Newspaper Pages for "Real Time Print Media Monitoring"

Overview

Partners: **zhaw**, **data lab**

The Project: Automated article segmentation of newspaper pages for real-time print media monitoring.

Most Successful Approach [3]: A 3D visualization of the data processing pipeline.

Combination: Integration of various data sources.

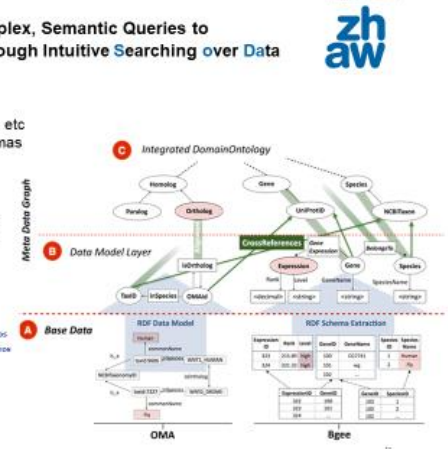
Result: Segmented newspaper articles.

Bio-SODA: Enabling Complex, Semantic Queries to Bioinformatics Databases through Intuitive Searching over Data

- Intuitive exploration
- without knowing SPARQL, SQL, etc
 - without knowing database schemas
 - large datasets

- Impact
- large bioinformatics user bases
 - future federation of life sciences

Lead: Kurt Stockinger, ZHAW



2. Research projects

Volume

- > 9 Mio. CHF 3rd party funding in first 4 years

Topics: all of digitalisation

- Industry 4.0 (e.g., CTI project QualitAI)
- E-Health (e.g. SystemsX/SNSF project AneuX)
- FinTech (e.g., CTI project DatFrisMo)
- Mobility (e.g., project Placebook)
- Sustainability (e.g., CTI project EAT-IT CO₂)
- Technology (e.g., SNSF project Bio-SODA)
- Document Analysis (e.g., CTI prj. DeepScore)
- ...

2. Research projects

Volume

- > 9 Mio. CHF 3rd party funding in first 4 years

Topics: all of digitalisation

- Industry 4.0 (e.g., CTI project QualitAI)
- E-Health (e.g. SystemsX/SNSF project AneuX)
- FinTech (e.g., CTI project DatFrisMo)
- Mobility (e.g., project Placebook)
- Sustainability (e.g., CTI project EAT-IT CO₂)
- Technology (e.g., SNSF project Bio-SODA)
- Document Analysis (e.g., CTI prj. DeepScore)
- ...

AneuX: Is aneurysm shape significant for its hazardousness?

Web tool for statistical analysis

ARGUS MEDIA BASED INTELLIGENCE

PANOPTES – Automated Article Segmentation of Newspaper Pages for "Real Time Print Media Monitoring"

Most Successful Approach [3]

Bio-SODA: Enabling Complex, Semantic Queries to Bioinformatics Databases through Intuitive Searching over Data

Intuitive exploration

- ✓ without knowing SPARQL, SQL, etc
- ✓ without knowing database schemas
- ✓ large datasets

Impact

- large bioinformatics user bases
- future federation of life sciences

Lead: Kurt Stockinger, ZHAW

Hydrobus: s

The challenge

- Not enough training

The project

- Self-adaption of control to time-varying demands in a multi-operation building using simulations
- Combined entropy and energy optimization of HVAC-system based on Model Predictive Control
- Integrates renewable energy technology, social dynamics and scenario-based weather prediction

The upside

- Enables a Swiss SME to harvest results from modern mathematics, data science and AI
- Gives science the opportunity to test modern approaches on real-world problems

Result

ies to hing over Data

75 Big Data

2. Research projects

Volume

- > 9 Mio. CHF 3rd party funding in first 4 years

Topics: all of digitalisation

- Industry 4.0 (e.g., CTI project QualitAI)
- E-Health (e.g. SystemsX/SNSF project AneuX)
- FinTech (e.g., CTI project DatFrisMo)
- Mobility (e.g., project Placebook)
- Sustainability (e.g., CTI project EAT-IT CO₂)
- Technology (e.g., SNSF project Bio-SODA)
- Document Analysis (e.g., CTI prj. DeepScore)
- ...

AneuX: Is aneurysm shape significant for its hazardousness?

Web tool for statistical analysis

ARGUS MEDIA BASED INTELLIGENCE

PANOPTES – Automated Article Segmentation of Newspaper Pages for "Real Time Print Media Monitoring"

Overview

Most Successful Approach [3]

SystemsX.ch fun

- Morphological analysis
- Biologically motivated
- Creation of datasets
- Creation of data
- Development of

Partners (co-apply)

Hydrobus: Simulation-based Optimization

Result

Hydrobus: s

The challenge

- Not enough training data for AI in socio-technological systems

The project

- Self-adaption of control to time-varying demands in a multi-apartment building using simulations
- Combined entropy and energy optimization of HVAC-system based on Model Predictive Control
- Integrates renewable energy technology, social dynamics and scenario-based weather prediction

The upside

- Enables a Swiss SME to harvest results from modern mathematics, data science and AI
- Gives science the opportunity to test modern approaches on real-world problems

RINO Electronics AG

ies to hing over Data

Base Data

2. Research projects

Volume

- > 9 Mio. CHF 3rd party funding in first 4 years

Topics: all of digitalisation

- Industry 4.0 (e.g., CTI project QualitAI)
- E-Health (e.g. SystemsX/SNSF project AneuX)
- FinTech (e.g., CTI project DatFrisMo)
- Mobility (e.g., project Placebook)
- Sustainability (e.g., CTI project EAT-IT CO₂)
- Technology (e.g., SNSF project Bio-SODA)
- Document Analysis (e.g., CTI prj. DeepScore)
- ...

AneuX: Is aneurysm shape significant for its hazardousness?

Aneurysm on x-ray image, Isolated aneurysm for shape analysis, Cells of vascular wall, Model of vascular wall, Web tool for statistical analysis

SystemsX.ch fun
 • Morphological a
 • Biologically moti
 • Creation of dise
 • Creation of data
 • Development of

Partners (co-appl
 Zürcher Partnerin
 AneuX Shape as I
 for Aneurysm

AneuX: Is aneurysm shape significant for its hazardousness?

Aneurysm on x-ray image, Isolated aneurysm for shape analysis, Cells of vascular wall, Model of vascular wall, Web tool for statistical analysis

SystemsX.ch funding: 2M CHF, review by SNSF
 • Morphological analysis of aneurysms using machine learning
 • Biologically motivated simulation model of cell wall change
 • Creation of disease model for treatment planning
 • Creation of database of aneurysms
 • Development of tools to analyze clinical (image) data

Partners (co-appl
 Zürcher Partnerin
 AneuX Shape as Biomarker for Aneurysm Disease
 Universität Zürich
 HUG
 UNIVERSITÉ DE GENÈVE
 ETH zürich
 FOUNDATION
 Swiss Neuro Foundation
 cobmm
 RINO Electronics AG

ARGUS MEDIA BASED INTELLIGENCE
PANOPTES – Automated Article Segmentation of Newspaper Pages for "Real Time Print Media Monitoring"

Overview, Most Successful Approach [3], Result

Hydrobus: s

Initial setting by engineer, CONTROL (e.g. AMB), Training with model data

The challenge
 • Not enough training

The project
 • Self-adaption of
 • Combined entropy and energy optimization of HVAC-system based on Model Predictive Control
 • Integrates renewable energy technology, social dynamics and scenario-based weather prediction

The upside
 • Enables a Swiss SME to harvest results from modern mathematics, data science and AI
 • Gives science the opportunity to test modern approaches on real-world problems

ies to hing over Data

Integrated DomainOntology

Big Data, Base Data, RDF Data Model, RDF Schema Extractor

zhaw, Unil (Université de Lausanne), Swiss Institute of Bioinformatics

2. Research projects

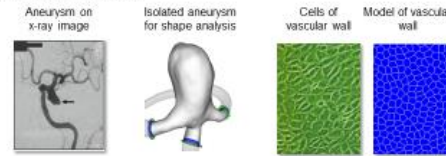
Volume

- > 9 Mio. CHF 3rd party funding in first 4 years

Topics: all of digitalisation

- Industry 4.0 (e.g., CTI project QualitAI)
- E-Health (e.g. SystemsX/SNSF project AneuX)
- FinTech (e.g., CTI project DatFrisMo)
- Mobility (e.g., project Placebook)
- Sustainability (e.g., CTI project EAT-IT CO₂)
- Technology (e.g., SNSF project Bio-SODA)
- Document Analysis (e.g., CTI prj. DeepScore)
- ...

AneuX: Is aneurysm shape significant for its hazardousness?

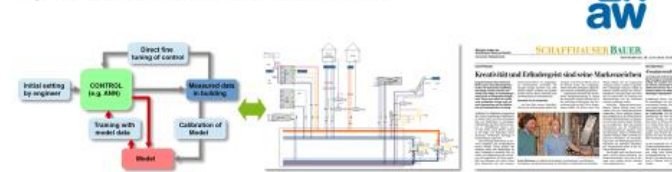


- SystemsX.ch funding: 2M CHF, review by SNSF
- Morphological analysis of aneurysms using machine learning
 - Biologically motivated simulation model of cell wall change
 - Creation of disease model for treatment planning
 - Creation of database of aneurysms
 - Development of tools to analyze clinical (image) data

Partners (co-applicant Sven Hirsch, ZHAW):



Hydrobus: Simulation-based Optimization



- The challenge
- Not enough training data for AI in socio-technological systems

- The project
- Self-adaption of control to time-varying demands in a multi-apartment building using simulations
 - Combined entropy and energy optimization of HVAC-system based on Model Predictive Control
 - Integrates renewable energy technology, social dynamics and scenario-based weather prediction

- The upside
- Enables a Swiss SME to harvest results from modern mathematics, data science and AI
 - Gives science the opportunity to test modern approaches on real-world problems



ARGUS MEDIA BASED INTELLIGENCE

PANOPTES – Automated Article Segmentation of Newspaper Pages for "Real Time Print Media Monitoring"

Most Successful Approach [3]

Overview

Partners: The Project

Combination

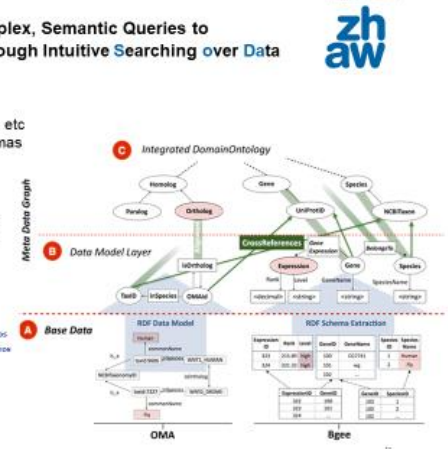
Result

Bio-SODA: Enabling Complex, Semantic Queries to Bioinformatics Databases through Intuitive Searching over Data

- Intuitive exploration
- without knowing SPARQL, SQL, etc
 - without knowing database schemas
 - large datasets

- Impact
- large bioinformatics user bases
 - future federation of life sciences

Lead: Kurt Stockinger, ZHAW



3. Education activities

Undergraduate

- Involved in numerous courses of B.Sc. Programs
 → e.g., «scripting», «big data», «data mining», «AI», «information retrieval», «data warehousing», ...

Graduate and post-graduate

- Master of Science in Engineering specialization** in data science: e.g., «machine learning»
- Ph.D. programs:** with Universities of Venice, Zurich, and Neuchatel



swissuniversities

Professional education

- Master of Advanced Studies (**MAS**) in **Data Science**
- Completely booked** almost until end of 2019

<p>CAS Data Product Design Data-specific Service Design, Data-specific Business Models, Practice workshop, Security & Privacy</p>	<p>CAS Machine Intelligence Machine Learning, Deep Learning, Text Analysis, Advanced topics in Big Data</p>	<p>CAS Statistical Modeling Information processing with R, Advanced regression modeling, Analysis of time to event data, Network analysis</p>
<p>CAS Information Engineering Scripting in Python, Information Retrieval & Text Analytics, Databases & SQL, Data Warehousing, Big Data</p>	<p>CAS Data Analytics Data Description & Visualization, Statistical Foundations of Analytics, Multiple Regression, Time Series & Forecasting, Clustering & Classification</p>	

4. Highlights



Community involvement: Started Swiss Conference series on Data Science

- Grew from 120 participants to 270 (2014→2017)
- Handed it over to Data+Service community in 2018 (400 participants)
- Full academic track with IEEE sponsorship planned for 2019 (June 14, Bern)



Generating impact

- Co-initiator of **National Thematic Network**: Swiss Alliance for Data-Intensive Services
- **Book**: Braschler et al. (Eds.): *“Applied Data Science”* (Springer 2018, to appear)

