

Strategic field of research 2017–2020

The ETH Domain launches the Initiative for Data Science in Switzerland

The ETH Domain launches the Initiative for Data Science in Switzerland to accelerate data science through both an expansion of education and research, and the provision of infrastructure for data science users across disciplines.

In the past few years, data science has become extremely important internationally, with the majority of top-tier international research and teaching institutions investing significantly in dedicated centers and programs. The Initiative for Data Science in Switzerland will foster collaboration with top institutions and will include exchange of best practices in curriculum development and joint data science research projects advancing the state of the art in data science.

Data science is a strategic field of research of the ETH Domain for the period 2017–2020 and the Initiative will ensure that the ETH Domain and Switzerland possess the necessary expertise and remain globally competitive.

Swiss Data Science Center

The Initiative for Data Science in Switzerland will create the Swiss Data Science Center (SDSC), which will address the fragmentation within today's data and analytics landscapes (see box). Led jointly by EPFL and ETH Zurich, the SDSC will be composed of a distributed multi-disciplinary team of data scientists and experts in domains including personalized health and personalized medicine, earth and environmental science, social science and digital humanities, and economics. For example, research in genome-based personalized medicine requires new, efficient and privacy-aware analysis of molecular and clinical data; this is a big data challenge and currently a major bottleneck.

The vision for SDSC is to unlock actionable insights from diverse sources of data. The center aims to federate data providers, data and computer scientists, and subject-matter experts around a cutting-edge analytics platform offering domain-specific "Insights-as-a-Service" while addressing security and privacy issues inherent to the field of data science. The unique synergy that the center will enable among the institutions of the ETH Domain and between academic and industrial stakeholders in both data science and across carefully selected domains is expected to foster scientific breakthroughs with significant societal impact.



Data Science – Discovering actionable insights from complex data sources. (Image: iStock.com/derrek)

To accomplish this vision, the center will:

- develop a network of embedded data science support to work closely with research groups and foster collaboration between users and data scientists;
- offer end-to-end data science services to the research and development communities in Switzerland and beyond, specifically a set of hosted software solutions to provide "Insights-as-a-Service";
- create a community to share tools, methods and knowledge in the field.

Rather than building its own infrastructure, the online services of the SDSC will be backed by existing infrastructures of the ETH Domain (in particular by leveraging resources at the Swiss National Supercomputing Centre CSCS), SWITCH, as well as those of cloud providers.

Data Science

The 4th industrial revolution is driven by the convergence of smart, connected systems with breakthroughs in areas ranging from gene sequencing to nanotechnologies. In this new era where "data is the new oil", there is little value in "crude" data. But if crude data can be extracted, refined, and piped to where it can impact decisions, its value will soar.

Data science is the new research paradigm concerned with executing this vision. It sits at the intersection of several academic disciplines including data management and engineering, statistics, machine learning, algorithms, optimization and visualization. It offers a new tool to social sciences, economics, medicine, environmental sciences, and others, to understand and influence complex, real-world systems to make progress on some of the most challenging problems of our time.

Unfortunately, as things stand today, the available data and analytics methods and systems are highly fragmented. They are difficult to use for non-experts and the results are often very difficult to interpret. In other words, there is a gap between those who create data, those who develop data analytics and systems, and those who could potentially extract value from it. This gap is deepened by the inherent sensitivity of many data: trust through legal compliance, privacy preservation, and secure storage and processing are essential to develop data science to its full potential.

The Initiative for Data Science in Switzerland aims to address these challenges for the benefit of Switzerland, the scientific community, and the world at large.

Master in Data Science

The rapid rise of data science has created the need for both thoroughly trained data science specialists with a new profile of expertise, and basic knowledge of data science methods for researchers and practitioners in a broad range of disciplines. The number of students enrolled in data science related courses increases rapidly. Responding to this demand and interest, master courses in data science at EPFL and ETH Zurich and a catalogue of service courses targeted towards groups of application areas will be set up. The latter may be extended later into a continued education program to provide training to Swiss professionals.

The goal of the new master courses in data science is to produce graduates with thorough knowledge in data science methods and topics ranging from mathematical foundations, analytics, to software, systems, and real-world applications of the methods.

The core part of the master curriculum will consist of thorough training in data analytics, including courses in machine learning, statistics, data mining, natural language, image and signal processing, and artificial intelligence. This core is complemented by courses in data management, cloud computing and computer systems, security and privacy, and programming models. This will enable the graduates to become leaders in their professions. EPFL and ETH Zurich both plan to offer these new master courses in data science as of fall 2017.

Further information

www.datascience.ch

Published by

ETH Board
Dr Fritz Schiesser
President
Haldeliweg 15, CH-8092 Zurich
fritz.schiesser@ethrat.ch
www.ethboard.ch

Enquiries

Swiss Data Science Center (SDSC)
Dr Olivier Verscheure
Executive Director
EPFL SDSC, INN 314, Station 14
CH-1015 Lausanne
olivier.verscheure@epfl.ch

Zurich and Berne, May 2016