DLCM
Research Data Management Hub

#SRDD2018

Swiss Research Data Day 2018
Programme

THE SWISS RESEARCH DATA LANDSCAPE: CHALLENGES AND OPPORTUNITIES

12 June 2018
ETH Zurich, Room Audi Max

There has been a data deluge in the research sector in recent years, accompanied by an increasing tendency to allow the scientific community access to original research data and the procedures used in order to enable study results to be fully reproduced. This increasingly complex research data landscape presents new challenges, but also new opportunities. A good understanding of pre-conditions of sharing data, of the life cycle, and of the patterns of creation, curation, use and re-use of research data are required. The management of research data is far from a straightforward issue, and clear strategies are required in order to ensure that data remain accessible and understandable over time. Furthermore, funding bodies, including the Swiss National Science Foundation (SNSF), are reacting to this changing landscape, with research data management plans now a key requirement for funding rather than an optional extra.

In addition, recent Data Science approaches are combining data from different sources to obtain novel findings. Society as a whole is also becoming increasingly involved in the scientific learning process through initiatives such as citizen science, whereby the general public can contribute data to the research process and draw their own conclusions. In the Swiss Research Data Day 2018, established experts across different domains will share best practices around handling, managing, and sharing research data.

DLCM | swissuniversities

Contents

Practical information

Room Audi Max (HG F 30)

ETH Zurich
Main building
Rämistrasse 101
8092 Zurich
Switzerland

Contact

info@dlcm.ch

Programme overview

09:00-10:00 ARRIVAL AND REGISTRATION

10:00-10:15 WELCOME AND INTRODUCTION

By Ulrich Weidmann, Vice President Personnel and Resources, ETHZ

10:15-11:00 OPENING KEYNOTE

Shared Infrastructures for Open Science, by Prof. Kilian Stoffel, Rector of the University of Neuchâtel and President of the Steering Committee of swissuniversities' Program "Scientific information" - P-5

DLCM project overview, by Dr. Pierre-Yves Burgi, DLCM project Director and Deputy CIO, UNIGE

Short break

11:15-12:00 BREAK-OUT SESSIONS

Data Management Plan (DMP): From Theory to Practice - DMP Templates, Best Practices & Guidelines	EnhanceR and DLCM – Solving Data Challenges for Swiss Researchers	Hands-on workshop: Creating cross-linked deposits in the Zenodo repository of CERN	Using Patient Data in Health Research: Challenges and Opportunities	The Human Brain Project Open Neurodata Ecosystem
Room AudiMax	*Room D1.2*	*Room D7.1*	*Room D7.2*	*Room D1.1*

Lunch and poster session

13:15-14:45 PRESENTATIONS

But can I trust your data?, by Dr. Marjan Grootveld, Senior policy officer at DANS

Cultural data sculpting in public and research domains, by Prof. Sarah Kenderdine, Professor of Digital Museology at EPFL

From liberating imprisoned data in publications to publishing linked open data, by Dr. Donat Agosti, Founding President of Plazi GmbH

14:45-15:30 BREAK-OUT SESSIONS

Data Management Plan (DMP): From Theory to Practice - DMP Templates, Best Practices & Guidelines	EnhanceR and DLCM – Solving Data Challenges for Swiss Researchers	Hands-on workshop: Creating cross-linked deposits in the Zenodo repository of CERN	Using Patient Data in Health Research: Challenges and Opportunities	A multidimensional approach to Open Science practices: illustrative examples from Swiss universities
Room D1.1	*Room D1.2*	*Room D7.1*	*Room D7.2*	*Room AudiMax*

Coffee break

16:00-17:00 PRESENTATIONS

Data to the People - The Power of Citizen Science in Healthcare, by Prof. Ernst Hafen, Professor of Systems Genetics at ETH Zurich and president of MIDATA.coop

Opening Science in a Data-Driven World, by Dr. Rok Roškar, Data Scientist / Software Engineer, Swiss Data Science Center

17:00-17:15 SUMMARY AND CONCLUSION

Detailed programme

PRESENTATIONS

10:15-10:45 **Shared Infrastructures for Open Science**
by Prof. Kilian Stoffel, Rector of the University of Neuchâtel and President of the Steering Committee of swissuniversities' Program "Scientific information" - P-5

'Scientific information: Access, processing and safeguarding' is a funding program managed by swissuniversities, the rectors' conference of Swiss higher education institutions. The program has been granting project-based contributions since 2013. Contributions intend to be an incentive for Swiss universities to build and share infrastructures for scientific information. The program, now in a consolidation phase, aims at building a network of nationally shared services in scientific information, supported by a coordinating organization, operating from 2021 onwards. The program is proud to fund key projects such as DLCM. The keynote will be an opportunity to sum up achievements and develop on the strategic preliminaries for 2021-2024, in the context of Open Science.

10:45-11:00 **DLCM project overview**
by Dr. Pierre-Yves Burgi, DLCM project Director and Deputy CIO, UNIGE

Launched officially in August 2015 and funded by swissuniversities (SUC P-2 / P-5 programs), the DLCM project aims to establish national services for the most important data management researchers' needs, such as: training, and personalized advices for writing DMPs; active data management solutions; long-term storage options and data publication according to international standards. Particular emphasis was also placed on the services' sustainability by applying business model methodologies. At the term of its first phase, this talk will present the main project outcomes and the national vision for the next years.

13:15-13:45 **But can I trust your data?**
by Dr. Marjan Groovfeld, Senior policy officer at DANS

Openness, sharing, FAIR data, citizen science – O brave new world, that has such vision in't! For some researchers this is no longer visionary, but already practice. For many others, however, notions like Open and FAIR data or data sharing are challenging. How can we make new data FAIR and how can we measure the FAIRness of existing data? "Trust" is an essential aspect: trust that others will interpret "your" data correctly, for instance, or trust in repository systems.

13:45-14:15 **Cultural data sculpting in public and research domains**
by Prof. Sarah Kenderdine, Professor of Digital Museology at EPFL

Within research institutions cultural and social data has become vital sources for new research, exemplified by work in the digital humanities. In the collecting institutions (galleries, libraries, archives and museums), digital technologies and datasets have the power to expand collecting policies, renew curatorial visions and refresh relationships with audiences. In this presentation, the opportunities for engaging the public with

tangible and intangible heritage data will be explored through a series large-scale interactive and immersive experiences that result from collaborative and interdisciplinary research practises. While the opportunities for creative reuse of (big and small) cultural data are immense, there are diverse challenges including infrastructure, access, data curation and copyright. These challenges require national strategic approaches that can support the future data-driven research and engagement.

14:15-14:45 **From liberating imprisoned data in publications to publishing linked open data**
by Dr. Donat Agosti, Founding President of Plazi GmbH

Scientific publications ought to contribute to the dissemination of scientific results. Whilst the quintessence of current scientific publications lays in creating unstructured publications from often highly structured data, the Swiss company Plazi does the reverse, creating structured, findable, citable, and reusable data from scientific publications. Since data is not work in a legal sense, and thus is not copyrighted the Plazi workflow can operate under Swiss copyright law – a competitive advantage for Switzerland. This workflow is based on standalone open source application written in Java with modular Natural Language Processing tools that can be chained into highly customizable pipelines allowing TDM to be highly automated including discovering from named entities to entire specific text blocks. This workflow is used to build two repositories to disseminate this data. TreatmentBank includes text based entities such as taxonomic treatments (220,000 extracted from 31,000 articles), close to 1M bibliographic references or observation records (80,000), and millions of facts such as geo-data, specimen data or scientific names. The Biodiversity Literature Repository (BLR), a community within CERN's Zenodo repository, includes a growing corpus of 172,000 extracted scholarly figures, and 32,000 article deposits with very rich metadata. Access to the articles itself remain closed if published as closed access. The upload is fully automated. Each of the deposits includes meta data, links to the source article, related items, and a digital object identifier (DataCite DOI). Currently, BLR is with 51% percent Zenodo's largest data provider, but at the same time occupies only 1.2% of the total storage space – an invitation to rapidly increase the amount of data made accessible. Furthermore, each of the extracted elements can be cited using globally unique persistent identifiers. This liberated data is upon upload immediately submitted to some of the world's largest science infrastructure, such as the US National Center for Biotechnology Information or the Global Biodiversity Information Facility. To avoid this complex task alltogether, Plazi is promoting the notion to publish scientific results as semantically enhanced publications in a collaboration with the Bulgarian Publisher Pensoft and the US National Library of Medicine. Currently over 20 inhouse and hosted journals are available, including Arpha, the needed advanced publishing platform to publish accordingly. Together, they provide services to convert and produce semantically enhanced publications starting from unstructured to upfront structured data and manuscripts.

16:00-16:30 **Data to the People - The Power of Citizen Science in Healthcare**

by Prof. Ernst Hafen, Professor of Systems Genetics at ETH Zurich and president of MIDATA.coop

Public health and medical research largely depends on the availability of personal data as diverse as genomic, geolocation, nutrition, fitness and medical data. The individual is only agent that is permitted to consent to the aggregation of these different data types. Personalised health research thus depends on the active participation of citizens/patients by collecting and contributing their data to research. For example, via dedicated smartphone apps patients can record sensor data and their wellbeing or side effects associated with new medications of surgical treatments. For such patient reported outcomes to be adopted widely, a new trust promoting framework for data sharing is needed. I will argue that citizen owned non-profit data cooperatives in which users remain in full control over their personal data provide such a framework. I will present results from several successful studies run on the MIDATA platform cooperative including a citizen science hay fever study which recruited over 7000 participants within the first week of launch of the Ally Science app. Since data can be copied citizen-controlled management of these data will not only improve healthcare it also offers the chance to democratise the personal data economy.

16:30-17:00 **Opening Science in a Data-Driven World**

by Dr. Rok Roškar, Data Scientist / Software Engineer, Swiss Data Science Center

Benefits from open science are numerous and various. The most obvious are transparency and verifiability. They are attributes that science requires from researchers who make a scientific claim public. Today, with the increasing volume and complexity of data, sharing scientific results in the form of an article is not sufficient to verify the validity of this claim. Open science in general promotes incentives, tools and best practices to share scientific results beyond the traditional publication format. In the light of the “reproducibility crisis” in various fields, it has become obvious to many scientists that transparency is one way to alleviate the risk of erroneous – or fraudulent – conclusions. Providing researchers with the skills and tools to properly document their studies will ultimately foster trust and excellence in science. Researchers must be able to trust published research even when the data is not made public. There is therefore a critical need for tools enabling a trusted, verifiable science at all times.

MORNING BREAK-OUT SESSIONS

11:15-12:00 **Data Management Plan (DMP): From Theory to Practice - DMP Templates, Best Practices & Guidelines**

by Dr. Aude Dieudé, EPFL Library, and Dr. Ana Sesartic, ETH Zurich

Did you recently prepare a DMP for your grant application? If so, you may already know that since 2017, public funding agencies such as the Swiss National Science Foundation (SNSF), now require detailed information on the life cycle of data when grant applications are submitted. As a result, at the stage of completing a grant application on mySNF, researchers must provide information regarding their Data Management Plan (DMP) to be eligible to receive funding. This DMP form comprises four sections: (1) data collection and documentation, (2) ethics, legal and security issues, (3) data storage and preservation, and (4) data sharing and reuse. Essentially, research data management is today a must-have skill rather than a nice-to-have option. However, what are the best practices, concrete guidelines and resources useful for all researchers and grant applicants? Once your DMP has been prepared and your research project selected by funding agencies, how can you best translate into reality what has been written in your DMP? In a complex environment like today's research landscape, the management of research data is far from a straightforward issue. As a consequence, collaborative and proactive strategies along with tailored approaches and tools are essential to ensure that data remain accessible, understandable and reproducible over time. During this workshop, concrete resources and tools will be presented in order to offer practical solutions to get started with the creation of a DMP and the lessons learned for ensuring there is no discrepancy between theory and practice.

11:15-12:00 **EnhanceR and DLCM – Solving Data Challenges for Swiss Researchers**

by Dr. Pierre-Yves Burgi and Hugues Cazeaux, UNIGE and Dr. Alex Upton, ETH Zurich

Research presents unique challenges in IT, particularly the hands-on processing, analysis, and management of research data. Supporting researchers in meeting these challenges has led to the creation of a network of specialist units at Swiss research organisations to provide Research IT support, allowing researchers to concentrate on their core tasks and accelerating time to results. The DLCM (www.dlcm.ch) and EnhanceR (www.EnhanceR.ch) projects are national initiatives to provide specialised research IT support to the entire Swiss research academic community, and achieve this goal by facilitating cooperation between academic Research IT groups in Switzerland. In this talk, an overview is given of the services that EnhanceR and DLCM offer to researchers for dealing with their research data. Common data challenges that researchers face are presented, along with examples of previous support projects from various research communities that have helped them overcome these. Finally, the support process is presented, helping researchers identify whether their research could benefit from specialist support.

11:15-12:00 **Hands-on workshop: Creating cross-linked deposits in the Zenodo repository of CERN**

by Dr. Tim Smith, CERN and Dr. Donat Agosti, Plazi GmbH

Data not only want to be free, but cited. Scholarly articles include a rich array of data, from named entities to entire subarticle elements that can be discovered and extracted in Text and Data mining. To make them available for discovery, reuseable and citable, they need to be deposited in a repository, semantically enhanced and provided with a persistent identifiers and link to the source document. A single upload of data from a scholarly article becomes quickly a complex task that is too time consuming for human upload. Plazi developed such a workflow that included over 200,000 deposits in CERN's repository Zenodo. Plazi and Zenodo will demonstrate in the hands-on how this upload operates. Based on a real example, the upload procedure will be described, the API explained and the resulting output, the enhanced article and the related images with all the embedded links shown.

11:15-12:00 **Using Patient Data in Health Research: Challenges and Opportunities**

by Prof. Dr. Torsten Schwede and Dr. Leila T. Alexander

The amount and complexity of health-related data which can be collected growing as breathtaking speed. This ranges from molecular constituents of cells (genomics and epigenetics, proteins, metabolites, etc.) to clinical phenotypes of diseases (imaging, electrophysiology, etc.) and to personal lifestyle and environment (lifestyle tracking, "quantified self"). Increasingly, wearable or embedded devices connected via a networks can provide real time or longitudinal data. Provided these "big health data" are accessible and used effectively for meaningful correlation and association studies, they have great potential for biomedical research and individual health care in the future. Increasingly, biomedical research may reverse the traditional paradigm "from experimental animals to humans" into "from first in humans to validation in animals". However, using patient and citizen data for research raises a series of technical, scientific, ethical and legal (e.g. data protection) questions which must be addresses carefully. In this session, we will discuss challenges and opportunities in using personal data for health research in the context of the Swiss Personalized Health Network SPHN.

11:15-12:00 **The Human Brain Project Open Neurodata Ecosystem**

by Jeff Muller, EPFL Campus Biotech

The vast majority of world-wide neuroscience research funding is spent on the collection of unique neuroscience datasets. These datasets are usually used in a single lab for a small set of papers and then archived, never to be used again. If a small fraction of this data could be reused for new science, it would represent a massive cost savings for funding agencies. While funders attempt to push in this strategic direction, Open Data is neither the norm nor is it well appreciated in the Neuroscience community. The Human Brain Project (HBP) is attempting to solve this in their HBP Platform development. While still under heavy development, the HBP is currently operating a collaborative, open neurodata ecosystem with an active, diverse user base. As HBP works towards the establishment of a European Infrastructure with strong Swiss leadership, the lessons learned and strategic choices provide insight on what approaches work in neuroscience and by extension, what approaches and tools might help to catalyse Open Data in other domains.

AFTERNOON BREAK-OUT SESSIONS

14:45-15:30 **Data Management Plan (DMP): From Theory to Practice - DMP Templates, Best Practices & Guidelines**

by Dr. Aude Dieudé, EPFL Library, and Dr. Ana Sesartic, ETH Zurich

Did you recently prepare a DMP for your grant application? If so, you may already know that since 2017, public funding agencies such as the Swiss National Science Foundation (SNSF), now require detailed information on the life cycle of data when grant applications are submitted. As a result, at the stage of completing a grant application on mySNF, researchers must provide information regarding their Data Management Plan (DMP) to be eligible to receive funding. This DMP form comprises four sections: (1) data collection and documentation, (2) ethics, legal and security issues, (3) data storage and preservation, and (4) data sharing and reuse. Essentially, research data management is today a must-have skill rather than a nice-to-have option. However, what are the best practices, concrete guidelines and resources useful for all researchers and grant applicants? Once your DMP has been prepared and your research project selected by funding agencies, how can you best translate into reality what has been written in your DMP? In a complex environment like today's research landscape, the management of research data is far from a straightforward issue. As a consequence, collaborative and proactive strategies along with tailored approaches and tools are essential to ensure that data remain accessible, understandable and reproducible over time. During this workshop, concrete resources and tools will be presented in order to offer practical solutions to get started with the creation of a DMP and the lessons learned for ensuring there is no discrepancy between theory and practice.

14:45-15:30 **EnhanceR and DLCM – Solving Data Challenges for Swiss Researchers**

by Dr. Pierre-Yves Burgi and Hugues Cazeaux, UNIGE and Dr. Alex Upton, ETH Zurich

Research presents unique challenges in IT, particularly the hands-on processing, analysis, and management of research data. Supporting researchers in meeting these challenges has led to the creation of a network of specialist units at Swiss research organisations to provide Research IT support, allowing researchers to concentrate on their core tasks and accelerating time to results. The DLCM (www.dlcm.ch) and EnhanceR (www.EnhanceR.ch) projects are national initiatives to provide specialised research IT support to the entire Swiss research academic community, and achieve this goal by facilitating cooperation between academic Research IT groups in Switzerland. In this talk, an overview is given of the services that EnhanceR and DLCM offer to researchers for dealing with their research data. Common data challenges that researchers face are presented, along with examples of previous support projects from various research communities that have helped them overcome these. Finally, the support process is presented, helping researchers identify whether their research could benefit from specialist support.

14:45-15:30 **Hands-on workshop: Creating cross-linked deposits in the Zenodo repository of CERN**

by Dr. Tim Smith, CERN and Dr. Donat Agosti, Plazi GmbH

Data not only want to be free, but cited. Scholarly articles include a rich array of data, from named entities to entire subarticle elements that can be discovered and extracted in Text and Data mining. To make them available for discovery, reuseable and citable, they need to be deposited in a repository, semantically enhanced and provided with a persistent identifiers and link to the source document. A single upload of data from a scholarly article becomes quickly a complex task that is too time consuming for human upload. Plazi developed such a workflow that included over 200,000 deposits in CERN's repository Zenodo. Plazi and Zenodo will demonstrate in the hands-on how this upload operates. Based on a real example, the upload procedure will be described, the API explained and the resulting output, the enhanced article and the related images with all the embedded links shown.

14:45-15:30 **Using Patient Data in Health Research: Challenges and Opportunities**

by Prof. Dr. Torsten Schwede and Dr. Leila T. Alexander

The amount and complexity of health-related data which can be collected growing as breathtaking speed. This ranges from molecular constituents of cells (genomics and epigenetics, proteins, metabolites, etc.) to clinical phenotypes of diseases (imaging, electrophysiology, etc.) and to personal lifestyle and environment (lifestyle tracking, "quantified self"). Increasingly, wearable or embedded devices connected via a networks can provide real time or longitudinal data. Provided these "big health data" are accessible and used effectively for meaningful correlation and association studies, they have great potential for biomedical research and individual health care in the future. Increasingly, biomedical research may reverse the traditional paradigm "from experimental animals to humans" into "from first in humans to validation in animals". However, using patient and citizen data for research raises a series of technical, scientific, ethical and legal (e.g. data protection) questions which must be addresses carefully. In this session, we will discuss challenges and opportunities in using personal data for health research in the context of the Swiss Personalized Health Network SPHN.

14:45-15:30 **A multidimensional approach to Open Science practices: illustrative examples from Swiss universities**

organized by Prof. Basma Makhlouf Shabou, HEG/HES-SO and André Hoffmann, UZH, with Dr. Christine Pirinoli, Vice-Rector Research & Innovation, HES-SO, Dr. Genevieve Joullie Dardier, Research Data Specialist, HES SO, and Dr. Sergio Maffioletti, EnhanceR project Director

The session will address the important question how to enable effective Open Science practices among researchers within a multidimensional approach. Researchers are not the only actors in science and knowledge production, they need a strategic framework on which they will align practices and choices and a legal framework to specify ethical behaviour without affecting the private sphere of individuals by data sharing. In addition, without infrastructure and technologies it will not be possible to implement the appropriate systems for opening up research. It is important to take into account all strategic, legal and technical dimensions.

1. Introduction (André Hoffmann, OpenAIRE)

2. Open Science Strategic Perspectives in the HES-SO Network Context (Christine Pirinoli, HES-SO)

Abstract: The speaker will contextualise the HES-SO Open Science view which is based on the dissemination, the exposure and the promotion of applied science outcomes using Open Science practices.

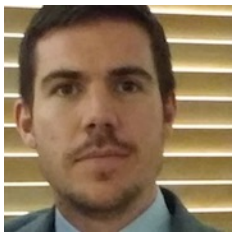
3. Multidimensional approach towards Open Science (Basma Makhlouf Shabou, HES-SO)

Abstract: This talk draws the main axes of the multidimensional approach enabling the Open Science practices in Swiss HEs. First, it attempts to evoke the legal and normative aspects of the different stages and tasks of the life cycle of research data management. In this respect, the basic principles of data processing and Archival Science and relevant initiatives will be highlighted. Second, methodological and ethical considerations and their role in guaranteeing the quality of shared data and results of scientific research will also be discussed with some applicable tools. The whole will be linked to the information culture which differs according to the type of institution, and/or the data, and/or the nature of the research disciplines.

4. IT for Open Science. Project EnhanceR - Enabling Open Science, Shareable, reproducible research containers (Sergio Mafioletti, EnhanceR)

Abstract: Reproducibility of data analysis is one of the cornerstones of Open Science, as data must be both available and reliable to support effective research. EnhanceR is a national initiative that develops and promotes services to help researchers making their research and results effectively reproducible. In this short presentation, we will show practices and tools to support this, such as orchestration of containers and validation of works-flows.

Speakers



Alex Upton

Community Project Manager, Scientific IT Services, ETH Zurich

After his doctorate in bioinformatics at the University of Birmingham, Alex completed a postdoctoral position at the Bitlab group at the University of Malaga. His research there involved extensive collaboration with local hospitals to analyse clinical data using computationally intensive methods to identify potential biomarkers of interest. In addition, he was involved in the proposal development and co-ordination of the training areas of the H2020 project Elixir EXCELERATE. He joined ETH Scientific IT Services in 2016 as Community Project Manager, and is involved in the Swiss national projects EnhanceR (www.EnhanceR.ch), and DLCM (www.DLCM.ch).

Ana Sesartic

ETH Zurich

Dr. Ana Sesartic is responsible for research data management training and consulting at the Digital Curation Office, ETH Library, ETH Zurich. As an environmental scientist with a PhD from ETH Zurich, she was confronted with the daily challenges of data management, especially in climate research. She is passionate about joining forces between disciplines in order to bridge the curation gap and make science more open and sustainable. During the Swiss national project dedicated to Data Life Cycle Management (DLCM), she was leading the ETH Zurich efforts regarding data management planning (DMP) and education. Thanks to a close collaboration across Swiss institutions and organisations, spearheaded by the EPFL and ETH Zurich, DMP templates and guidelines were established, which helped numerous researchers at all career levels to manage their research and conform to funders' regulations.

André Hoffmann

University of Zurich

André Hoffmann is long-term preservation coordinator, repository manager and Open Science advocate at the Main Library of the University of Zurich since 2012. As a member of the OpenAIRE project he serves as Open Science supporter for EU-funded projects in Switzerland. André Hoffmann holds a M.A. degree in sociology, political science and media science from the University of Konstanz, Germany.

Aude Dieudé
EPFL Library

Dr. Aude Dieudé is a project leader at EPFL since 2014. In close collaboration with her colleagues, Aude leads the personalized Data Management Plan (DMP) support service at the Rolex Learning Center, which is the first such tailored support service created in Switzerland. In addition, she offers tailored workshops and trainings on how to optimize Research Data Management (RDM) & DMP to leaders and researchers within Switzerland and abroad. On a national level, Aude is heading the first track of the Swiss national project dedicated to Data Life Cycle Management (DLCM), which offers RDM tailored tools, training and solutions. Thanks to a fruitful collaboration across Swiss institutions and organizations, her team provided sustainable resources and tangible solutions regarding DMP best practices, guidelines and policies for Switzerland. Aude holds a PhD from Duke University and has worked within academia, U.N. organizations and NGOs in the United States, Japan, and Switzerland. Passionate about bridging skills between different sets of expertise, she welcomes collaborating with a wide range of professionals and experts across fields as well as sectors such as academia, start-ups, private and nonprofit organizations, to name a few.

Basma Makhlouf Shabou
Professor of Information Sciences at Information Science Department, Geneva School of Business Administration, HES-SO

Dr. Basma Makhlouf Shabou has been Professor of Archival Science at the Geneva School of Business Administration (HEG) of University of Applied Sciences and Arts Western Switzerland (HES SO) since September 2010. She holds a Master's degree in Social Studies and a Postgraduate degree in Records Management followed by a doctoral research that she conducted at the University of Montreal in School of Library and Information Sciences (EBSI) in 2010. She is also involved in several international organizations, such as the Portail International Archivistique Francophone (PIAF) and the International Council on Archives (ICA), as a member of the Programme Commission (PCOM) and the Expert Group on Appraisal (EGA). She is active as researcher in several international research projects (InterPARES, RiC). Her research, lectures and publications focus on archival appraisal, information quality measurement, access and accessibility to public data, and the issue of information governance. She is also interested in research data management and is responsible for the training, consulting and teaching part of the Data Life Cycle Management (DLCM) project.

Christine Pirinoli
Vice-Rector Research & Innovation, HES-SO

Christine Pirinoli is Vice-Rector Research and Innovation at University of Applied Sciences and Arts Western Switzerland (HES SO). Doctor in social sciences, she has been previously working at the University of Lausanne as researcher and, for the latest 11 years, she has been both Professor and Director of research at the School of health sciences HESAV / HES-SO.

Donat Agosti
Founding President of Plazi GmbH

Donat Agosti is a biologist with a PhD degree from ETHZ. He had research positions at the Natural History Museum London, University of Zurich, American Museum of Natural History New York and the Jet Propulsion Laboratory, California Institute of Technology, Pasadena. He published over 90 scientific publications, including a widely cited manual on monitoring biodiversity. Involved since 1996 in fostering open access to biodiversity data, developing data conversion workflows, building two open data repositories, the Biodiversity Literature Repository at Zenodo and TreatmentBank. He was instrumental in setting up the Bouchout Declaration on Open Biodiversity Knowledge Management. Since 2008 he is founding president of Plazi, a Swiss company promoting open access to scientific results. He is a member of the Swiss University Council SUK P-5 expert panel.

Ernst Hafen
Professor of Systems Genetics at ETH Zurich and president of MIDATA.coop

Ernst Hafen, PhD, is a Professor of Systems Genetics at ETH Zurich and former President of ETH. In addition to over 30 years of academic research, he has founded and advised several biotechnology companies. He is the president of the BIO-TECHNOPARK Schlieren-Zurich. Ernst Hafen endeavors to assist scientific discovery and its efficient translation into products that help society and the economy. As a trained geneticist, he has a strong interest in human genetics and personalized medicine. He posits, that an individual's control over his or her personal health data, will be a key asset for better and more effective health care. In 2012 he acted as a founding member of the Association Data and Health and is the president of MIDATA.coop (https://www.midata.coop) which he co-founded in 2015. Citizen-owned personal data cooperatives enable citizens to securely store, manage and control access to their personal data and form the basis for a fair and sustainable personal data economy.

Genevieve Joullie Dardier
Research Data Specialist, HES SO

Genevieve Joullie Dardier is Research Data Specialist at the University of Applied Sciences and Arts Western Switzerland (HES SO). Prior to this appointment, she was Assistant Professor in Ocean Engineering at Stevens Institute of Technology (Hoboken, USA) and Senior Data Engineer for a private energy company. Her research and teaching interests include research data, ocean and glacier observation systems and community resilience. She is involved in the training and consulting part of the Data Life Cycle Management (DLCM) project as well as with Open Science issues at the HES-SO. She holds a B.S and a M.S degree in Physics from the EPFL and a PhD in Telecommunications from the National Telecom Institute in Paris, Telecom-Paristech.

Hugues Cazeaux
Head of e-Research, UNIGE

After an engineering degree in Computer Science, Hugues worked in different software companies during the last 20 years. Hugues used to introduce and to use agile methodologies in developing software. In the last decade, Hugues acquired a good expertise in archiving & record management domain. Today, Hugues is participating to DLCM project (ww.dlcm.ch) to design and build the solution to preserve Swiss research data.

Jeff Muller
EPFL Campus Biotech

Jeff became a software developer and entrepreneur working in the areas of sonar imaging, computer vision and highly interactive web applications for over 20 years. He has held senior management and executive positions in the game development and logistics software industries. In 2011, he joined the Blue Brain Project at École polytechnique fédérale de Lausanne to develop web-accessible, reproducible scientific workflows. In 2015 Jeff joined the Human Brain Project (HBP) central Project Coordination Office in Geneva, Switzerland. In the HBP, Jeff currently leads HBP Technical Coordination and the work package for Community-Driven Neuroinformatics Platform and Infrastructure Operations. He manages the teams which deliver the HBP Collaboratory, a web-based collaborative research platform. He also manages the team which builds the HBP Knowledge Graph, an open data publishing portal for neuroscience.

Kilian Stoffel

Rector of the University of Neuchâtel and President of the Steering Committee of swissuniversities' Program "Scientific information" - P-5

Kilian Stoffel became the rector of the University of Neuchâtel in August 2016. He studied mathematics and physics in Fribourg followed by a doctorate in computer science. He then worked as a researcher at the University of Maryland, the Johns Hopkins Hospital and University and the Hebrew University of Jerusalem. In 1997, Kilian Stoffel joined the University of Neuchâtel as a professor at the Institute of information management. He currently leads the steering committee of the Swiss P-5 program "Scientific information: access, processing and safeguarding".

Leila T. Alexander

Chief Operating Officer, SPHN Data Coordination Centre

Leila T. Alexander is the Chief Operating Officer at the SPHN Data Coordination Centre. She graduated as a Medical Engineer in 2009, earned her doctorate in drug discovery from the University of Oxford in 2013 and gained post-doctoral experience in cancer metabolomics at ETH Zürich. Leila has over 8 years' experience in bridging the gap between industry and academia, having worked at Novartis, AstraZeneca, a software development startup and health economics consultancy. She joined SIB Swiss Institute of Bioinformatics in 2016, where she is responsible for the Data Coordination Centre team capacity building and providing day-to-day operational oversight.

Marjan Grootveld

Senior policy officer at DANS

Marjan Grootveld is senior policy consultant at DANS in The Netherlands. DANS promotes permanent access to digital research information and encourages researchers to archive and reuse data. Marjan has been involved in OpenAIRE data management support for some years and has reviewed data management plans for the European Commission. She is also one of the coaches of the training “Essentials 4 Data Support”, which DANS and partners offer under the umbrella of Research Data Netherlands. “There is no progress without reuse”.

Pierre-Yves Burgi
DLCM project Director and Deputy CIO, UNIGE

Pierre-Yves Burgi is Deputy CIO, heading the unit for development and integration of IT solutions within the Information System Division of the University of Geneva. He is particularly involved in the e-research and e-learning domains. He is the initiator and director of the DLCM project which started in June 2014.

Rok Roškar
Data Scientist / Software Engineer, Swiss Data Science Center

Rok Roškar is an astrophysicist with a strong interest in distributed computation and data analysis. For the past several years he has been working on (big) data analysis problems in various domains within ETH. He is presently at the Swiss Data Science Center striving to make data science more efficient, transparent, reproducible, and repeatable.

Sarah Kenderdine
Professor of Digital Museology at EPFL

Professor Sarah Kenderdine researches at the forefront of interactive and immersive experiences for galleries, libraries, archives and museums. In widely exhibited installation works, she has amalgamated cultural heritage with new media art practice, especially in the realms of interactive cinema, augmented reality and embodied narrative. She is considered a pioneer in the field digital heritage, digital museology, digital humanities and data visualisation and is a regular keynote speaker at related forums internationally. In addition to her exhibition work she conceives and designs large-scale immersive visualisation systems for public audiences, industry and researchers. Since 1991 Sarah had authored numerous scholarly articles and six books. She has produced 80 exhibitions and installations for museums worldwide including a museum complex in India and received a number of major international awards for this work. In 2017, Sarah was appointed Professor of Digital Museology at the École polytechnique fédérale de Lausanne (EPFL), Switzerland where she has built a new laboratory for experimental museology (eM+), exploring the convergence of aesthetic practice, visual analytics and cultural data. She is also director and lead curator of EPFL's new art/science initiative, inaugurated in 2016 as ArtLab.

Sergio Maffioletti
EnhanceR project Director

After his doctorate in computer science at the University of Fribourg, Sergio worked at the Swiss National Supercomputing Center where he developed competences in High Performance and Grid computing. He joined the University of Zurich in 2009 as project director of the Grid Computing Competence Center. He was project manager of several initiatives for establishing a national academic e-infrastructure. He is now an infrastructure and application specialists at S3IT - the Service and Support for ScienceIT unit at the University of Zurich. There he consults and supports research groups in migrating complex data analysis usecases on large scale computational infrastructure; he helps institutes and research groups to understand their research infrastructure needs and how to develop a strategy for infrastructure services. Thanks to a dynamic and ever challenging environment, he developed competences spawning from infrastructure engineering, software development, project management, customer relationship management, but also as research consultant for a wide spectrum of research domains: from the scientific sector (chemistry, system biology, botany, geography, physics, biochemistry) to the financial one (economy, business and finance) to digital humanities (psychology, pedagogy, linguistics, history, archeology). He is also the project director of EnhanceR: a national initiative to provide specialised research IT support to the Swiss research academic sector (https://www.enhancer.ch/).

Tim Smith
CERN

Head of Collaboration, Devices and Applications Group at CERN, the European Particle Physics Laboratory. Tim is an Open Science advocate leading initiatives at CERN and in the wider science community. He drove the launch of CERN's Open Data Portal to share LHC big data with the world, as well as the Higgs Boson webcast which shared its discovery live around the globe. He also instigated and nurtures Zenodo within the European Commission's OpenAIRE project as an open data service for world-wide science.

Torsten Schwede
SIB Swiss Institute of Bioinformatics & University of Basel

Torsten Schwede acts as chair of the executive board of the SPHN Swiss Personalized Health network and is leading the SPHN Data Coordination Center activities. Torsten Schwede obtained his PhD on proteincrystallography in Freiburg i.Br., Germany, and then joined GlaxoSmithKline in Geneva as a bioinformatics scientist. In 2001 he accepted a position as professor for Bioinformatics at the Biozentrum of the University of Basel, and in 2002 became group leader at the SIB Swiss Institute of Bioinformatics. As scientific director of sciCORE he is responsible for the scientific computing infrastructure at the University of Basel.