HTW Chur
Swiss Institute for
Information Research

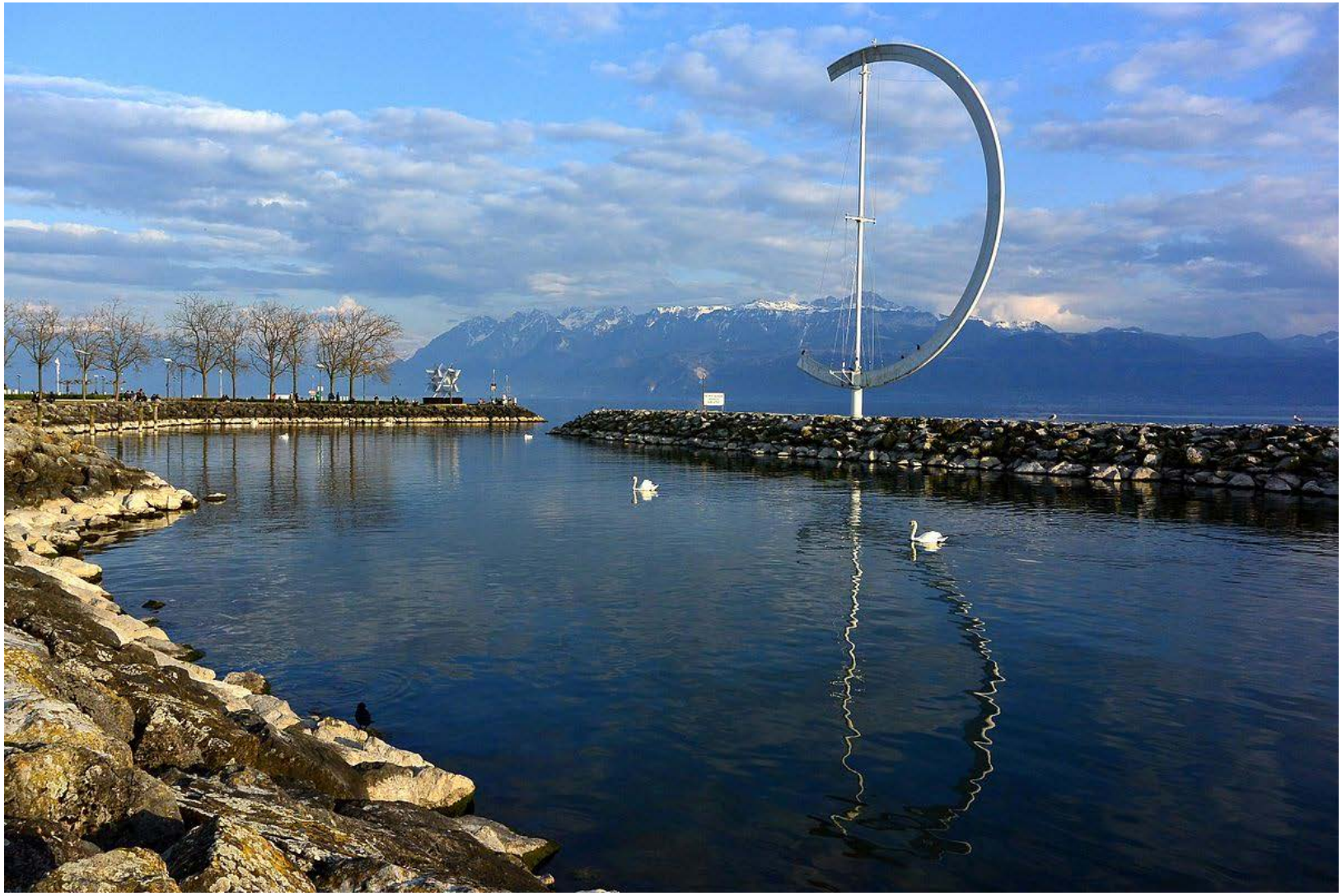
Research Data Management

A Swiss perspective

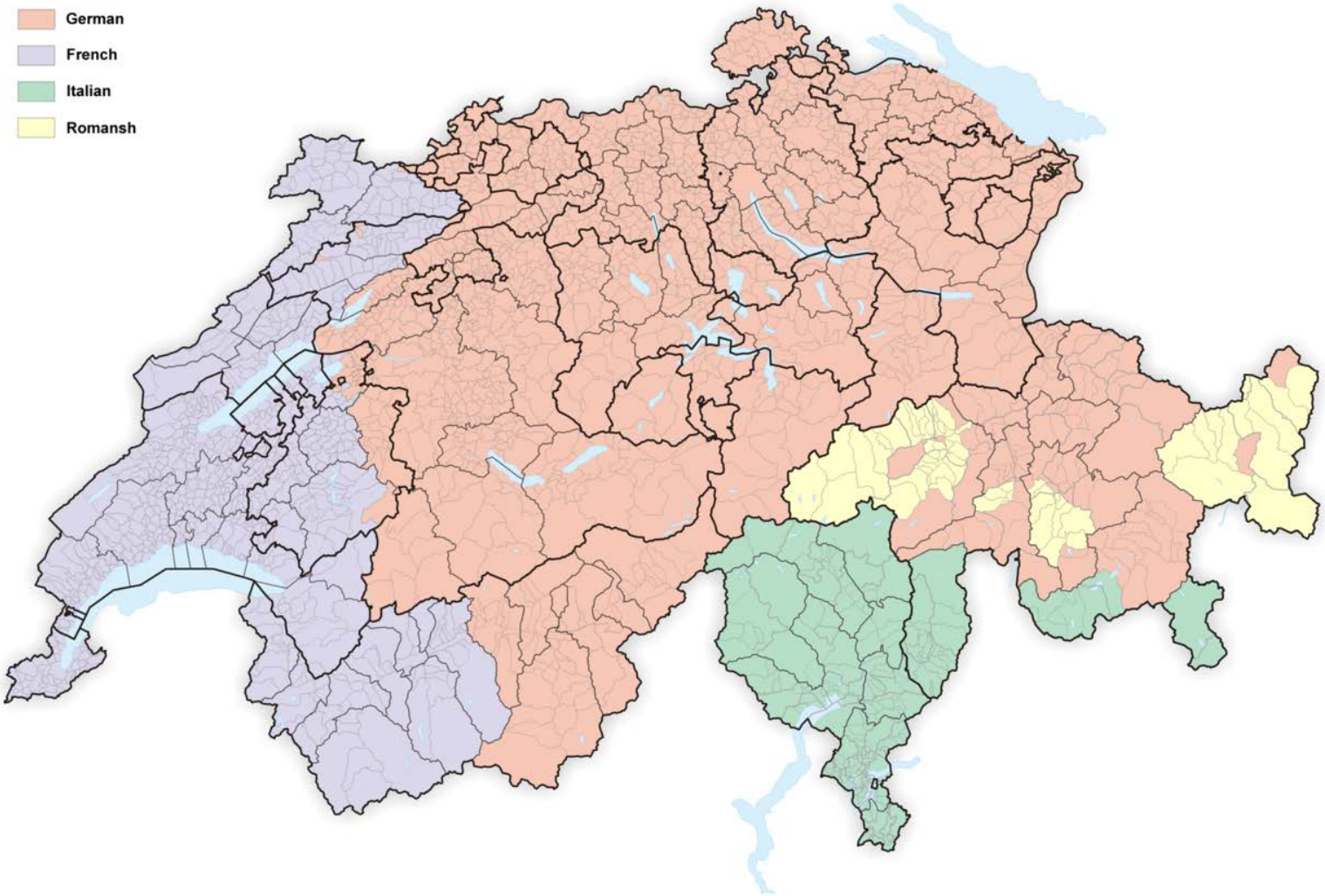
Prof. Dr. Ingo Barkow, Senior Lecturer for Operational Data Management

A Swiss perspective…..

Another Swiss perspective……



Switzerland in general



Research Data Management in Switzerland

Four levels / dimensions / perspectives:

- Political / Strategical level
- Infrastructure level
- Software Development
- Use of standards

Political / Strategical dimension

- High popularity of research data management as a topic with international and national policy makers
- Different approaches for organizing institutions
  - Centralized RDC approach (e.g. UK Data Service)
  - De-centralized RDC approach (e.g. German Data Council)
  - Library approach (e.g. US, „data librarian“)
  - Universities with institutional repositories
- Most countries have a mix of these approaches but nevertheless a dominating structure can be seen

Political / Strategical dimension

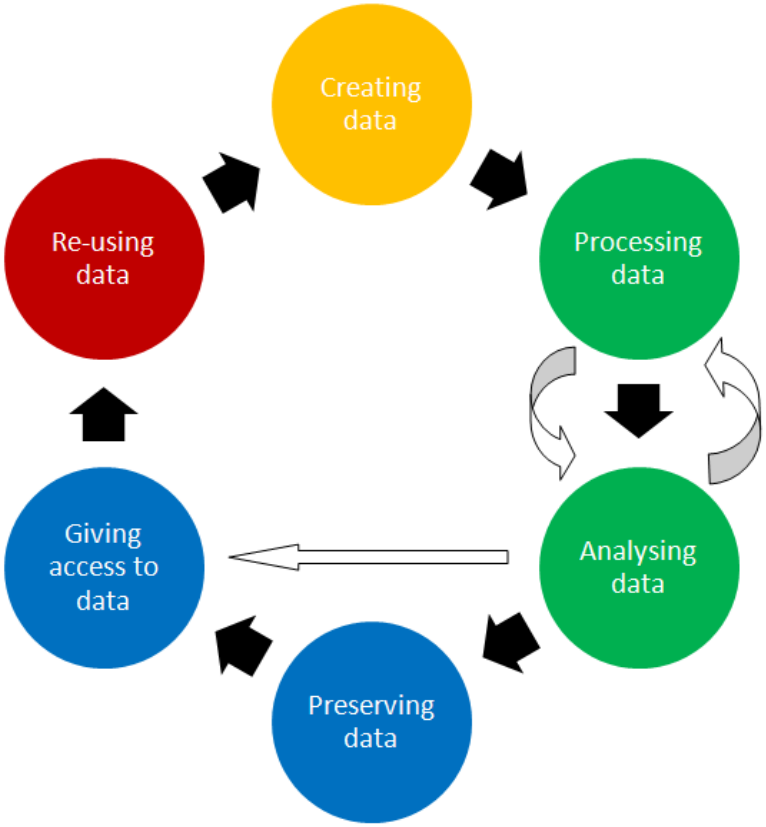
- Some players in Swiss Research Data Management
  - FORS (Switzerland's only Research Data Center)
  - University of Basel (Digital Humanities Lab)
  - CERN Library
  - ETH Library
  - Central Library Luzern (LORY Project)
  - RERO Library Foundation
  - HTW Chur (MScBA Program – Information and Data Science)
- Still unclear in Switzerland – RDM in federal RDC infrastructure or hosted at libraries?

Infrastructure dimension

- Policy makers see the need for data infrastructures
  - Research proposals need to outline their data management planning (e.g. EU Horizon 2020, Swiss National Fund)
  - Several calls for research data infrastructure
  - Larger projects (e.g. NEPS) establish their own data infrastructures
- Idea of common infrastructure projects to establish synergies between different organizations
- Re-use of data across country borders
- Enable open science on world-wide or at least European level

- Infrastructure projects are currently on demand in Switzerland as well (e.g. Swissuniversities SUK P-2 or SUK P-5 funding)

DLCM Project (started 2015)



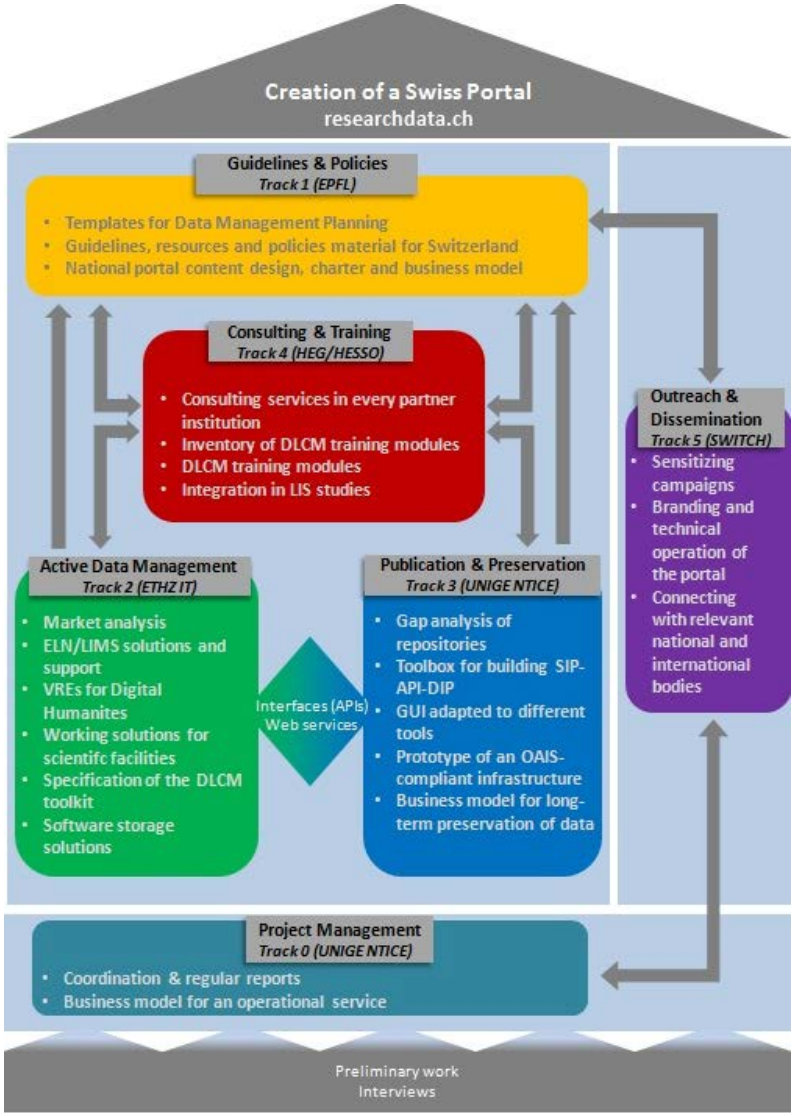
Digital Lifecycle Management Projekt (DLCM)

Eight partners:
EPFL, HEG, UNIL, UNIBAS, UNIZH, ETHZ, UNIGE und SWITCH

Project Homepage:
http://www.dlcm.ch

Kick-Off: 01.09.2015 (two years)

DLCM Project (started 2015)



Goals:

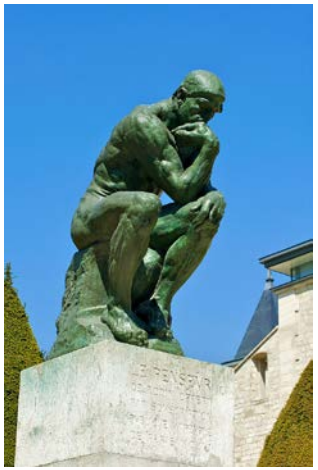
- Data Management Plan Template
- Guidelines
- Policy Template for Higher Education Institutions
- National portal
- Toolbox for building SIP, AIP, and DIP OAIS packages
- Prototype of a scalable OAIS-compliant infrastructure
- Business models for the delivery of viable long-term preservation services
- Inventory of existing data management training modules
- Specific data management training modules

Not all good infrastructure ideas work out in the end….

Data without Boundaries – DwB

The Data without Boundaries – DwB – project came to a formal end on 30 April 2015. The project had a mission to support equal and easy access to the rich resources of official microdata for the European Research Area, within a structured framework where responsibilities and liability would be equally shared. During its four-year lifespan the DwB worked towards preparing a comprehensive European service with better and friendly metadata, a more harmonized transnational accreditation and a secure infrastructure that would allow transnational access to the highly detailed and confidential microdata, both national and European, so that the European Union would be able to continuously produce cutting-edge research and reliable policy evaluations.

Challenges of large infrastructure projects



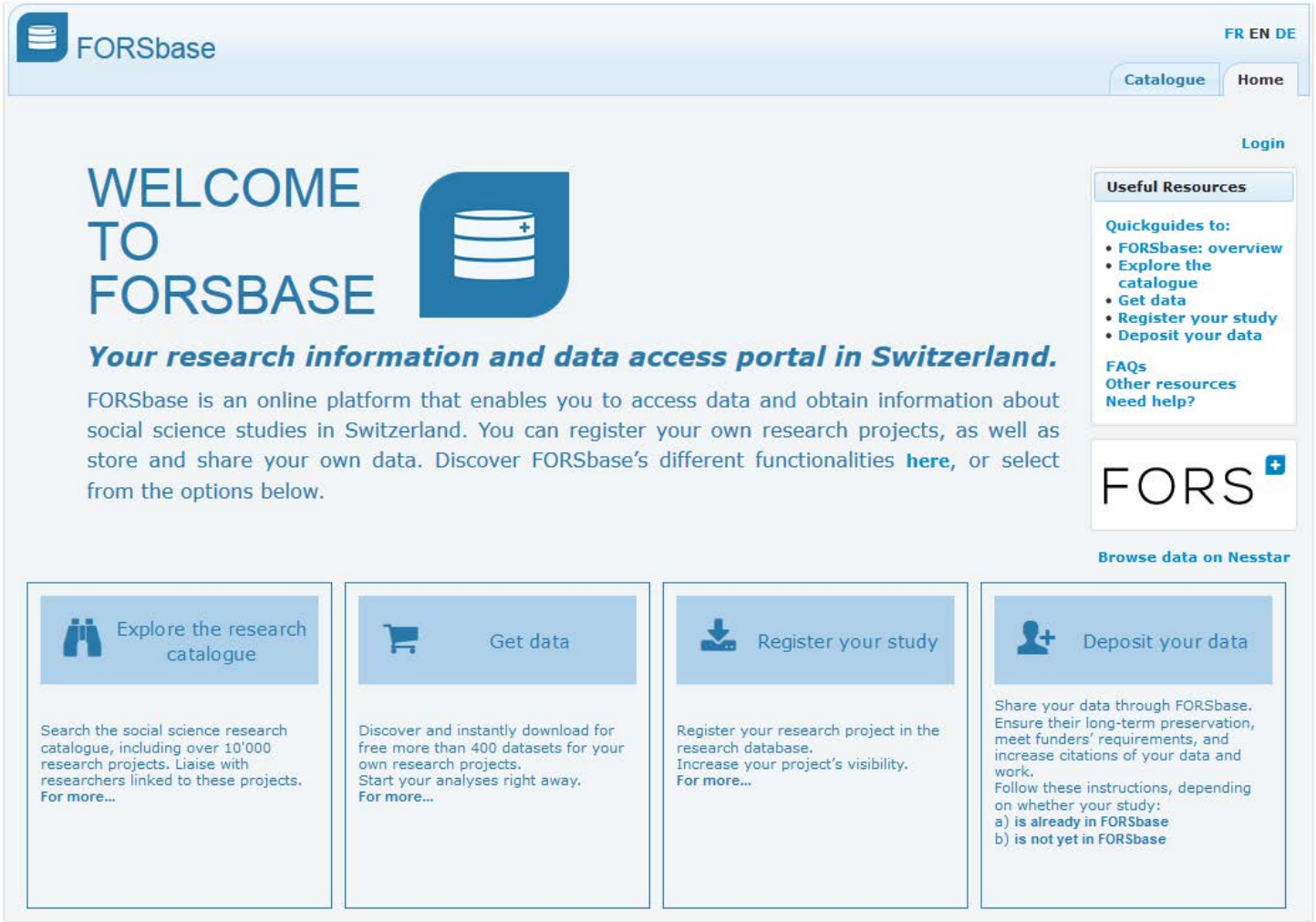
- Very often too ambitious
- Trying to be a «jack-of-all-trades» or «Swiss Army Knife»
- Too many partners at the very beginning
- Many university partners, only few implementation partners
- Handled like a research project
  - Project leadership by researchers / professors
  - Focus on production of concepts / papers
  - Implementation work shifted to second phase (which might never happen)
  - Large budget in total but underfinanced for the individual organization (0.5 – 1 research assistant for 2-4 years)
  - No sustainability (financing of infrastructure after building up is unclear)

Software Development

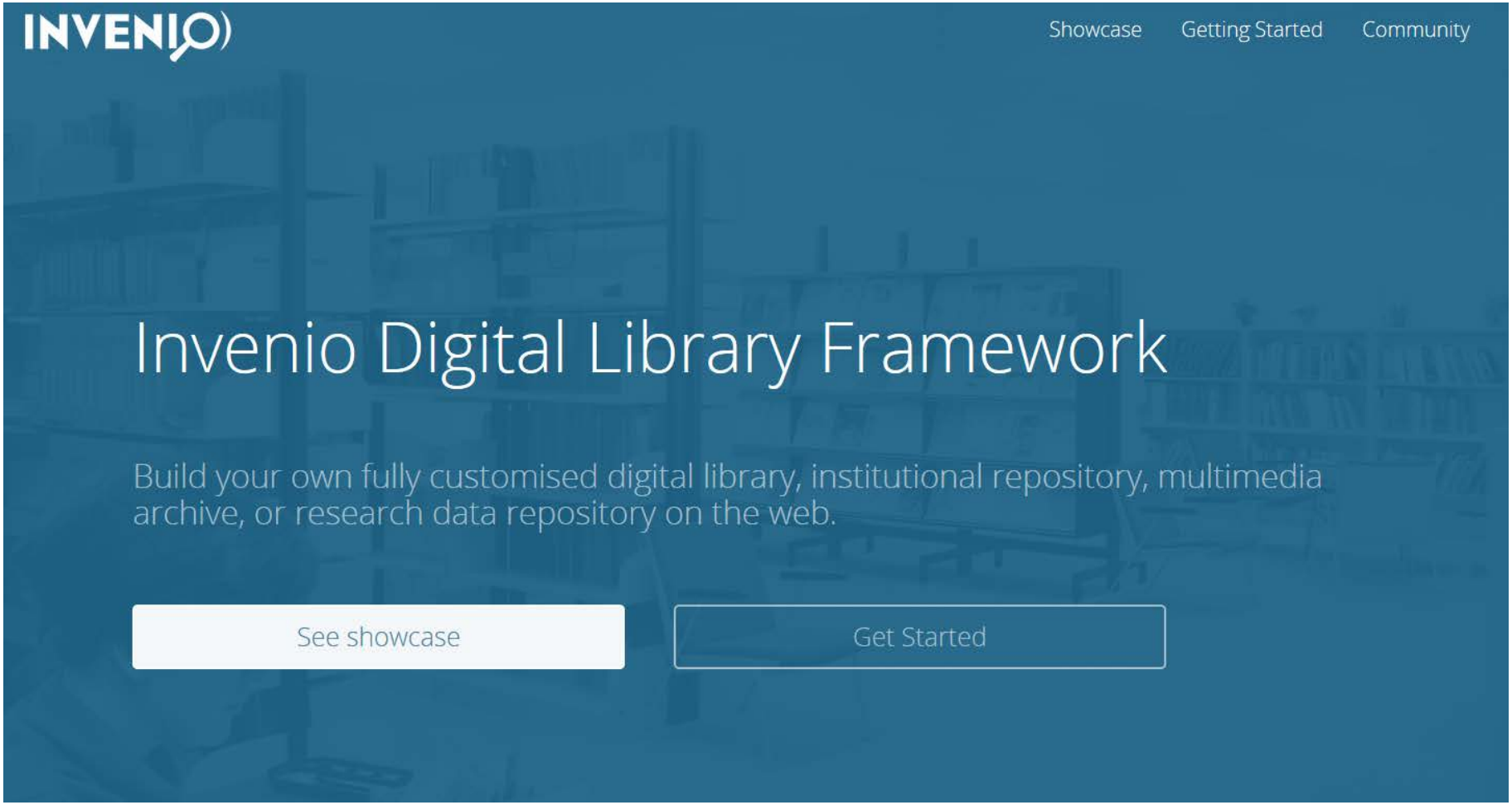
- Several software developments are triggered by organizational or infrastructure needs
- Some international examples
  - CESSDA
  - Colectica
  - Nesstar
  - RAIRD
  - Portage
  - Dataverse
  - Fedora

- Also a question for Switzerland – should we develop our own software or wait until an international package is available?

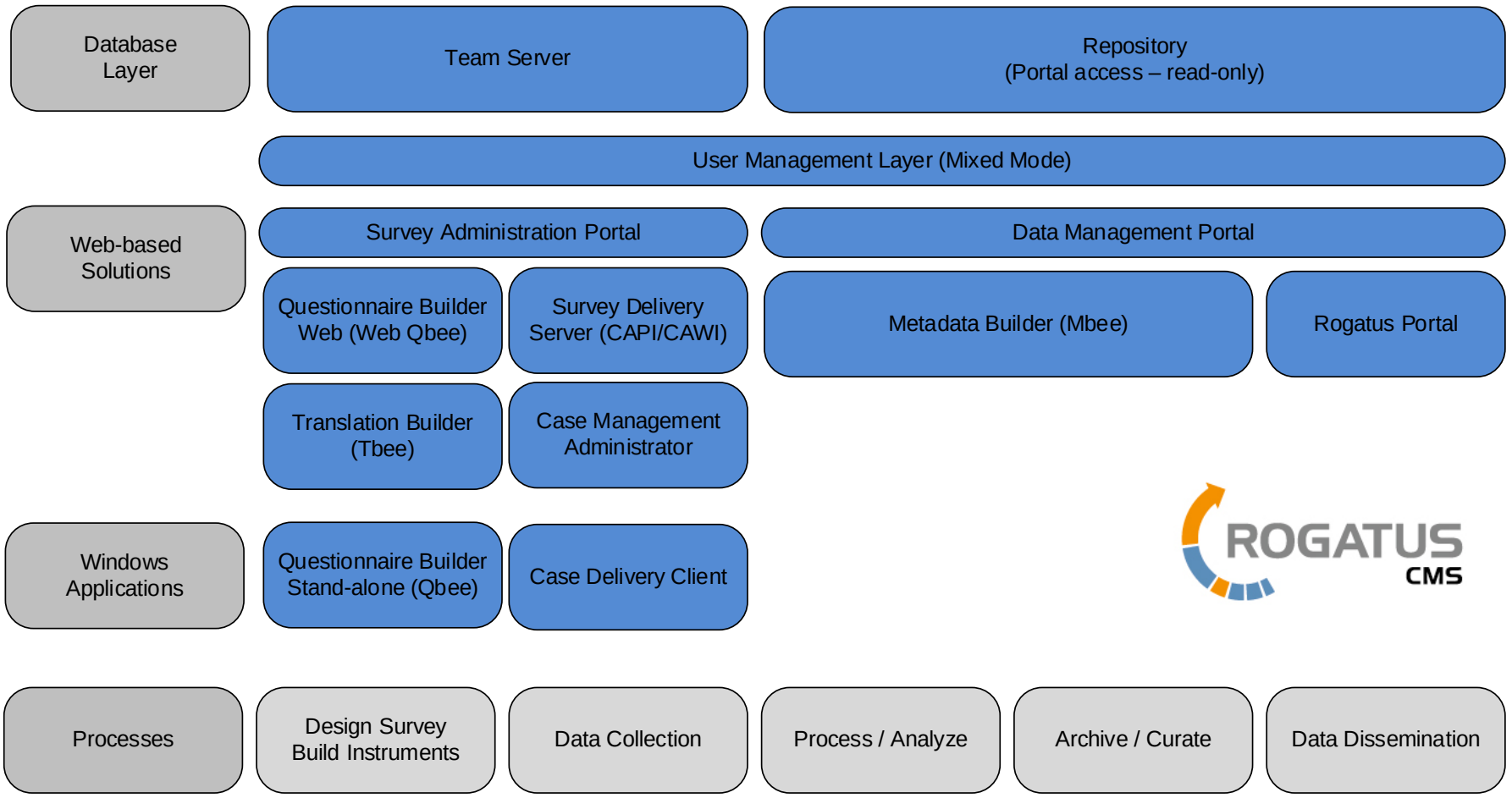
Swiss Repository Software - FORSbase



Swiss Repository Software – Invenio (CERN, tind)



R.I.P. – „Rogatus in Peace“ (DIPF/TBA21 2012 – 2015)



Software Development

- Very similar problems like infrastructure projects
- Many abandoned software projects (e.g. repositories, metadata editors)
- Especially lack of involvement of bigger commercial players
  - Maybe incentive is missing?
  - Maybe implementation especially in combination with standards is too complex?
- Software development driven by large surveys often results in “project-only software” or proprietary documentation standards

Use of standards

- Use of standards is often very uncomfortable for researchers
- Standards therefore have to be part of the software
- Invisible to users (“by-product of research process”)
- Some standards used in Switzerland
    - DDI (FORSbase)
    - MARC21 (Invenio)
    - JSON LD (Invenio)
    - OAI/PMH (Invenio)
    - Datacite Metadata Scheme
- Currently much more focus on metadata describing file formats for long term preservation, less focus on content metadata
- No focus on paradata or data collection

Software and standards – what does this mean for DDI?

- Support the development of DDI related software
- Publish a gap analysis focusing on needed software tools
- Guidelines for software usability in applications using DDI
- Support software developers through training and letters of commitment to funding agencies
- Create validation tools and profiles for DDI-C and DDI-L to assure interoperability of metadata created by different tools and organizations
- “Borrowed from SDMX” -> develop basic interfaces, web services and code examples for DDI implementations

In summary…..

- Switzerland needs a political debate how the overall research data infrastructure should be structured
- Infrastructure projects should have a pragmatic size and organizational fitting
- Software development should be embedded into an overall framework with individual modules developed in a pragmatic manner
- Standards should be community-driven and developed close to their technical implementation (“no tools, no standard”)

- Maybe re-using an idea from another community makes sense……

Agile Manifesto

“We are uncovering better ways of developing software by doing it and helping others do it. Through this work we have come to value:

Individuals and interactions over processes and tools
Working software over comprehensive documentation
Customer collaboration over contract negotiation
Responding to change over following a plan

That is, while there is value in the items on the right, we value the items on the left more.”

(Kent et. al. 2001 - http://agilemanifesto.org/)

„Agile Research Data Management Manifesto“

“We are uncovering better ways of handling research data management by doing it and helping others do it. Through this work we have come to value:

Processes and tools over academic debates and discussions
Working software over comprehensive documentation
Pragmatism over fine grained specifications
Use of standards over tackling individual research project needs

That is, while there is value in the items on the right, we value the items on the left more.”

Thanks for your attention.