

DLCM - The Swiss Way

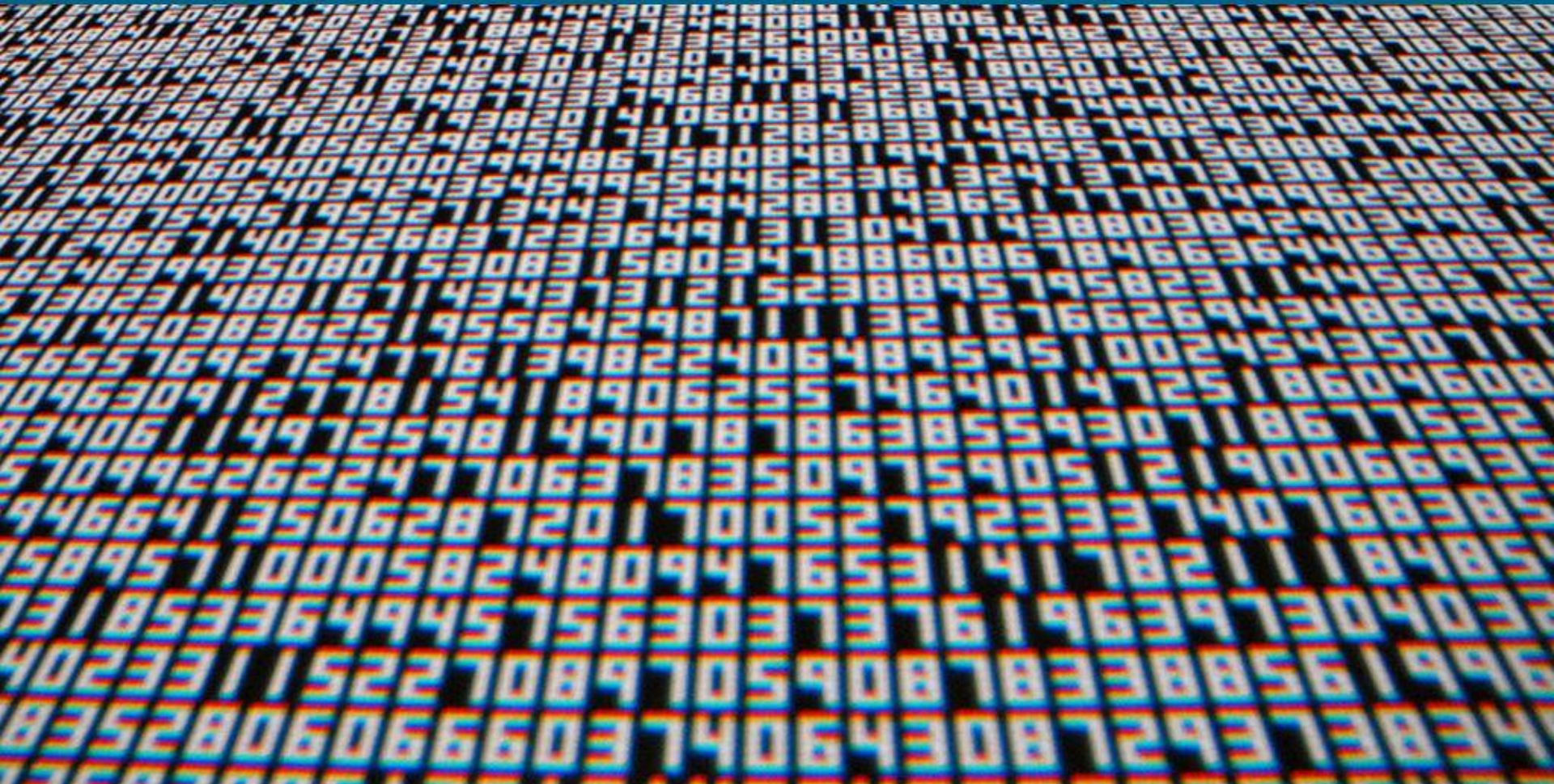
OAI9 - P.-Y. Burgi - Université de Genève



Agenda

- Swiss Context (SUC P-2)
- The DLCM Project
- Complex Objects?
- DLCM Tools
- Challenging Issues

Program SUC 2013-2016 P-2 "Scientific Information: Access, Processing and Safeguarding"



SUC-P2 Facts

- 36 M CHF

• 2013-2016 => extended

- 4 key areas:
Publications, eScience,
Basis, Services

Direction
de programme

Groupe
de soutien

Comité
d'experts

Architecture

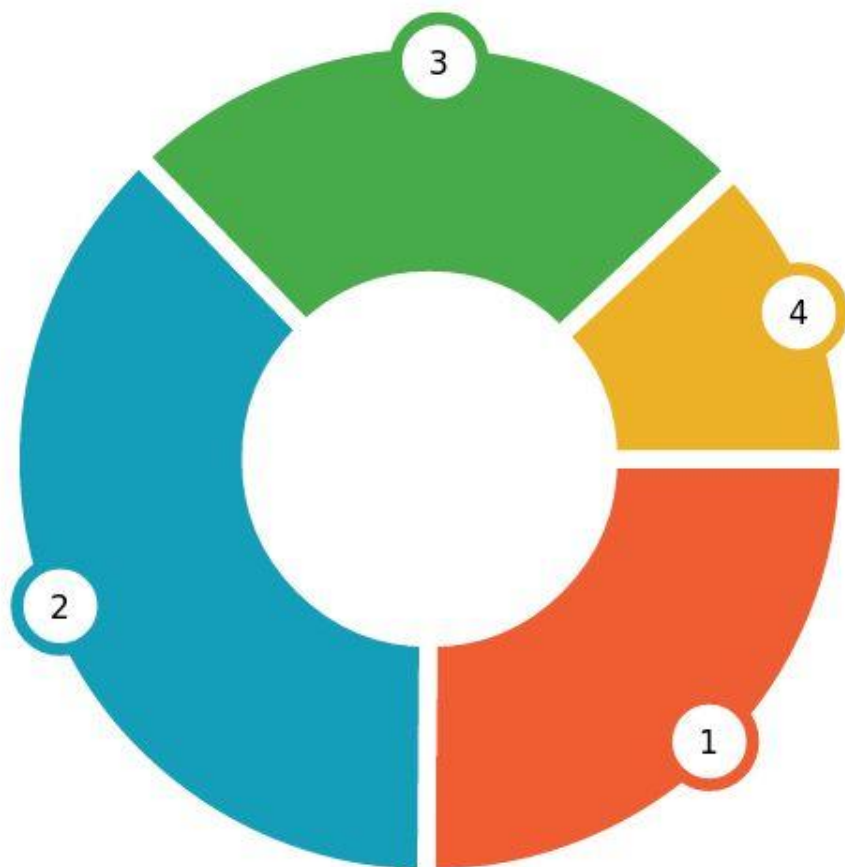
Direction
de projet

Opérateur
de service
(exploitation)

Projet de
mise en œuvre 1

Projet de
mise en œuvre 2

Currently 16 Active Projects

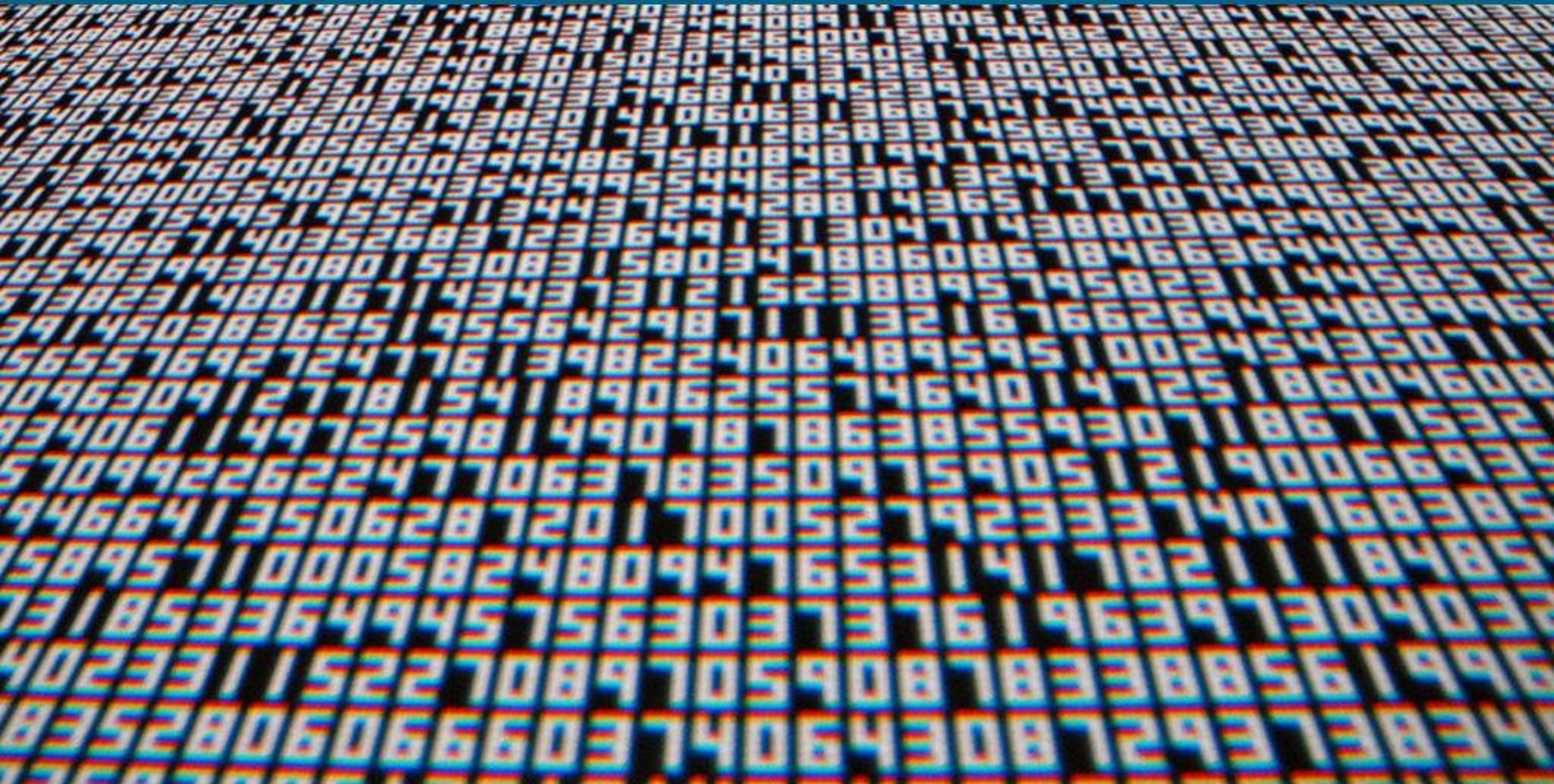


①	eScience	25%
②	Publications	38%
③	Basis	25%
④	Services	12%

eScience: Scope

- Developing concepts leading to national services
- Further developing established local services
- Supporting pilot projects
- Supporting training

The DLCM Project



8 Partner Institutions



ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE



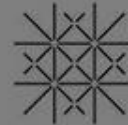
University of
Zurich ^{UZH}

Unil
UNIL | Université de Lausanne



UNIVERSITÉ
DE GENÈVE

h e g
Haute école de gestion
Genève



UNI
BASEL

SWITCH

ETH Zürich

Due to start 01.09.15
(final decision 22.06.15)





Pre-study phase

August 2014 to February 2015

RE/PRODUCE

will you be able to meet your

D A T A N E E D S

in the future?

how do you ensure your

R E S U L T S

are reproducible

could you benefit from

W I D E R A C C E S S

to your data

is the data safely

A R C H I V E D

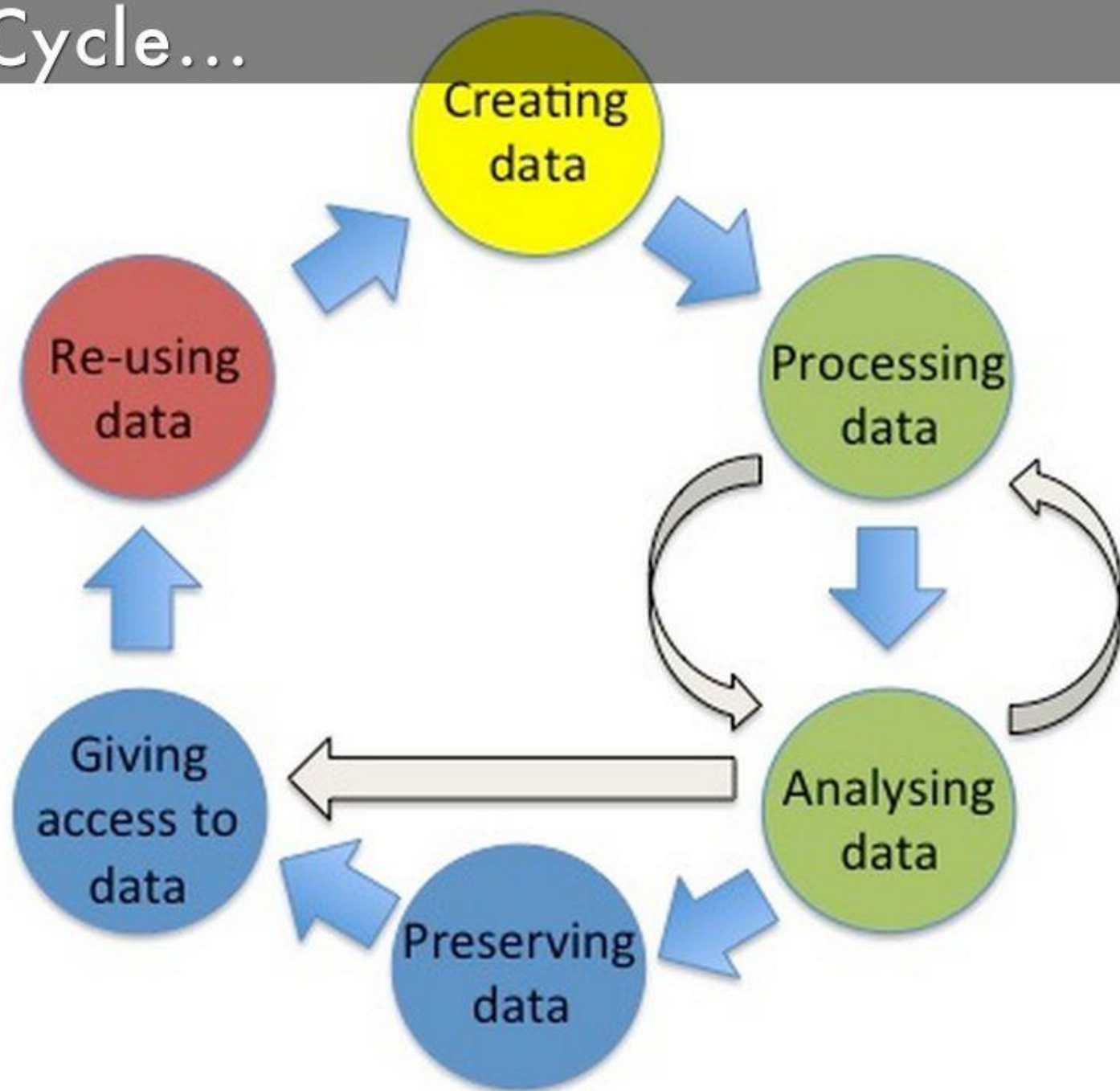
PUBLISH

The 4 key DLCM questions...

Interviews

Institution	Number of interviews	Disciplines
UNIGE	9	Topology, Mathematics, Linguistics, German, Cognitive Neuroscience, Educational Sciences, Geomatics, Archeology, Vulnerability, Political Sciences, Medicine (Child Cancer Research), Social Sciences
ETHZ	8	Biosystems Science and Engineering, Seismic Networks, Sociology, Consumer Behavior, Quantum Optics Group, Scientific Computing/ Photon Science, Physics
UNIL	15	Social Medicine, Social Sciences, Digital Humanities, Genomics, System Biology, Bio-informatics, Public Health, Imaging and Media Lab, Cancer Research
EPFL	5	Transport and Mobility, Quantum Optoelectronics, Supramolecular Nanomaterials and Interfaces, Audiovisual Communication Laboratory, Virology and Genetics.
UNIBAS	7	Biology research (Biozentrum, 2), Biology (Core facilities, 2), Molecular Psychology, Public Health (STPH), Digital Humanities
UNIZH	5	Law Science, Biology/Microscopy, Biology/Proteomics, University Hospital, Geosciences
Total	49	

"The" Cycle...



Swiss DM Portal

DLCM Project

Track 1
(EPFL)

Guidelines & Policies
(DMP, PMP, etc.)

Track 4
(HEG)

Consulting & Training

Track 2
(ETHZ)

Active
Data Management
(ELN/LIMS, VRE,
DLCM Toolkit, etc.)

APIs

Track 3
(UNIGE)

Publication &
Preservation
(OAIS, DOI, etc.)

Track 5
(SWITCH)

Outreach &
Dissemination
(national &
international
bodies)

DLCM Project

Links With Other CUS-P2 Projects

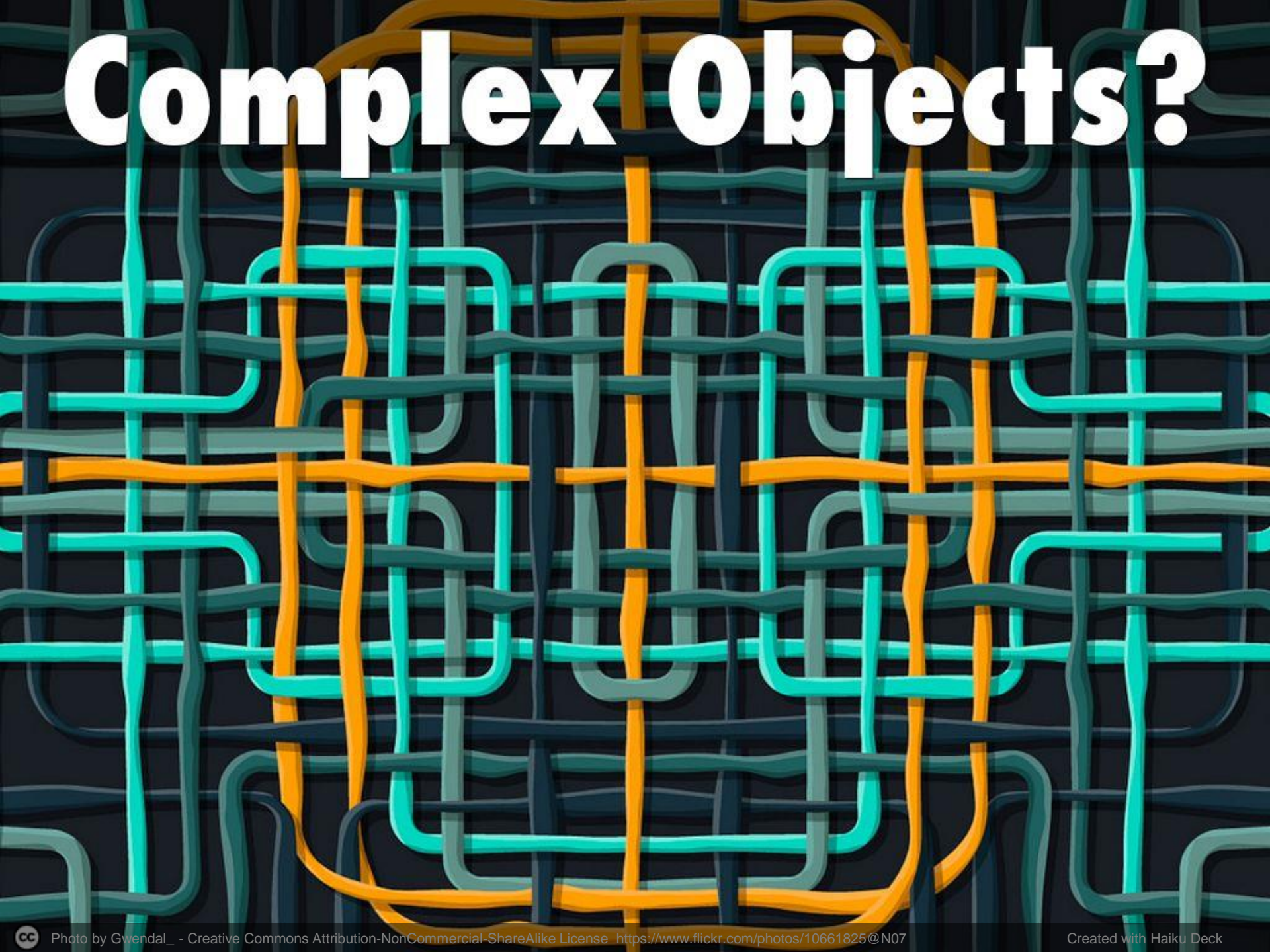


- eScience Coordination Team (eSCT)
- DICE+
- Open Research Data (ORD@CH)
- DACAR@HEG
- Data and Service Center for the Humanities (DaSCH)
- Swiss edu-ID and SCALE

Links With International Projects

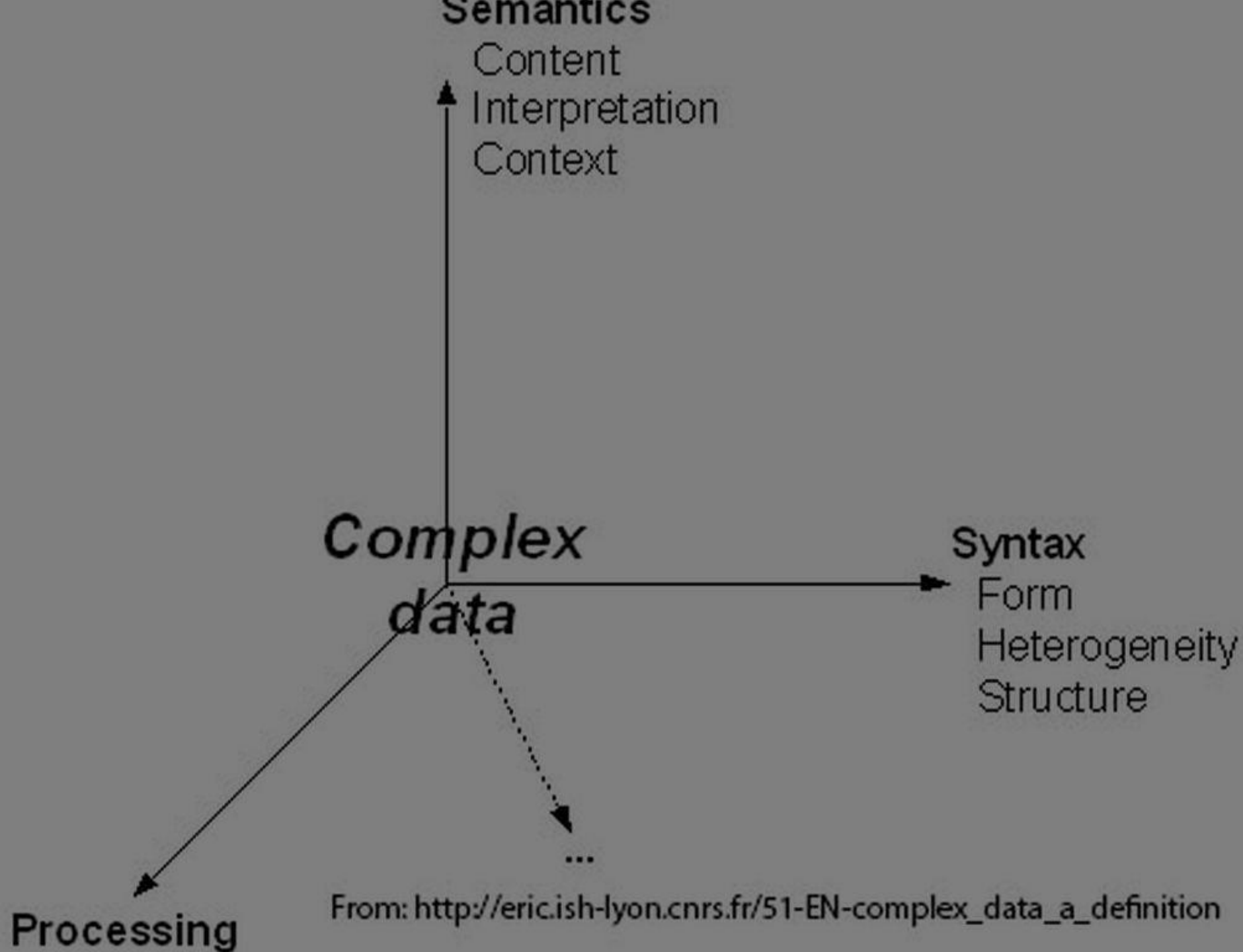
- Alliance Permanent Access
- OpenAIRE
- FAIRDOM: Findable, Accessible, Interoperable and Reusable (in biology domain)
- DARIAH

Complex Objects?



Simple Object

- Discrete digital files:
textual, pictorial, audio,
etc.



Complex Data

- Multiformat (DB, texts, images, sounds, videos...)
- Multistructure (relational databases, XML documents repository...)
- Multisource (distributed DB, Web...)
- Multimodal
- Multiversion (temporal DB...)
- Volumetry is contributing to complexity

DLCM Data, e.g.

- RDF databases (DH, genomic,...)
- Genome-based healthcare data of +30K individuals
- +40K scanned books (Bodmerlab)
- Imaging data in life-sciences

Bodmerlab (DH)

- World literature (160'000 works in eighty languages)
- High resolution images (6562 × 4452) - one per page
- +40K scanned books (out of 160K)
- Several hundreds of TBytes

Life Sciences

- Quantitative high-speed imaging of entire developing embryos with simultaneous multiview microscopy
- 175 million voxels per second for up to several days
- Several dozens of terabytes per specimen



The Tools

to manage research data

1. ELN / LIMS

- openBIS (@ETH-Z)
- SLims (Simple LIMS @ EPF-L)

MUST HAVE:
• NUMBERED PAGES
• TABLE INDEX

TODAY:

2 HOURS

PER DAY

MANAGING
RECORDING DATA

LAB NOTES

DIGITAL INVENTORY

LAB TOOLS - APS

SEARCHABLE

SHARED

BACKED UP

CREATES METADATA SETS

700K GRADS

500K INDUSTRIAL

MARKETING:

W-O-M WITH RESEARCHERS

TESTING WITH TWO LAB MANAGERS CURRENTLY

CLOUD-BASED

CONTENT CAN BE PRINTED AS PDF

(FOR HARD COPY)

\$5 DOWNLOAD + 90 DAY FREE → SUBSCRIPTION

evernote

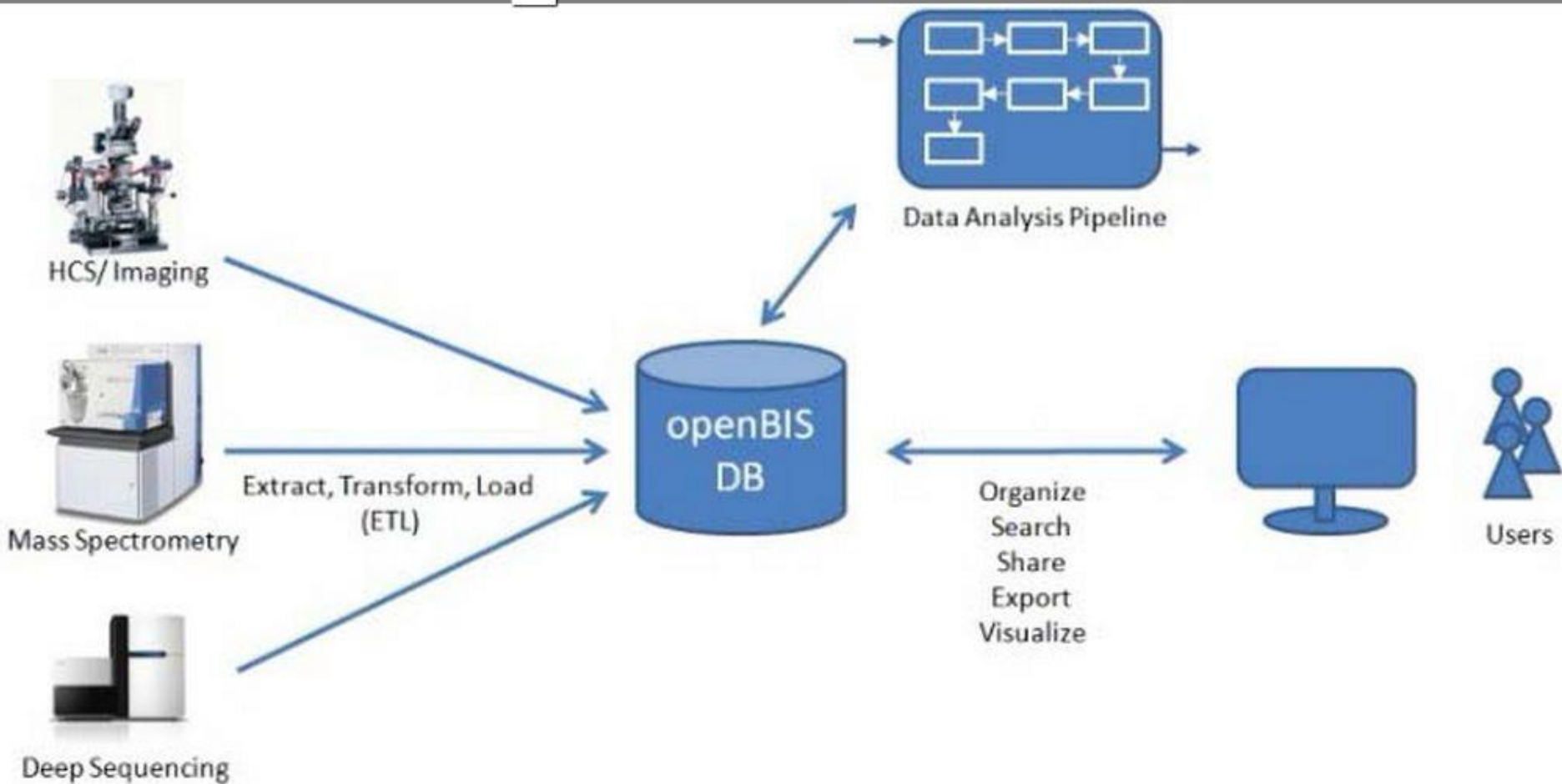
e-notebook
quantity

INVENTORY
MANAGEMENT

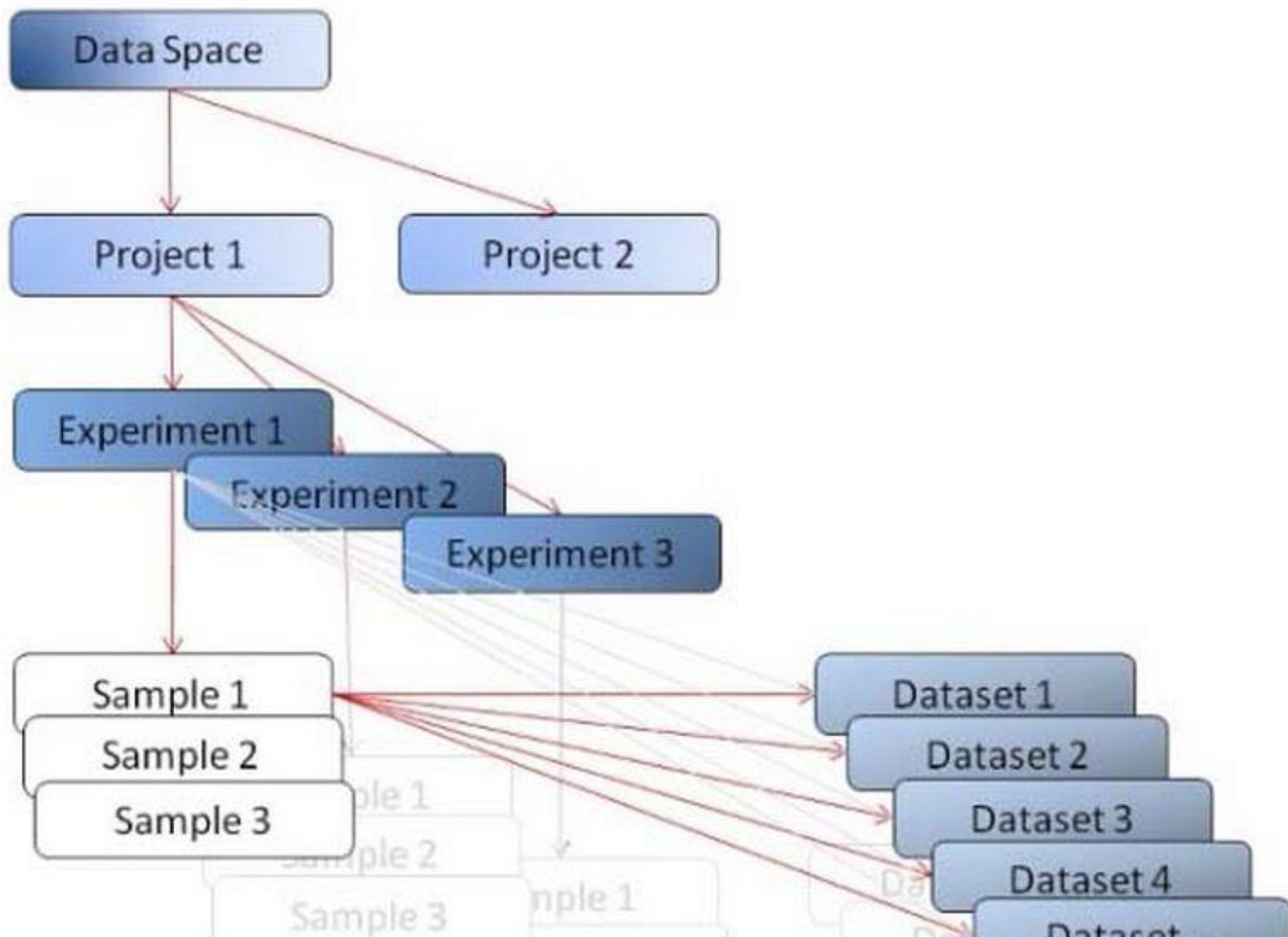
CAN BE RUN BY
ONE PERSON

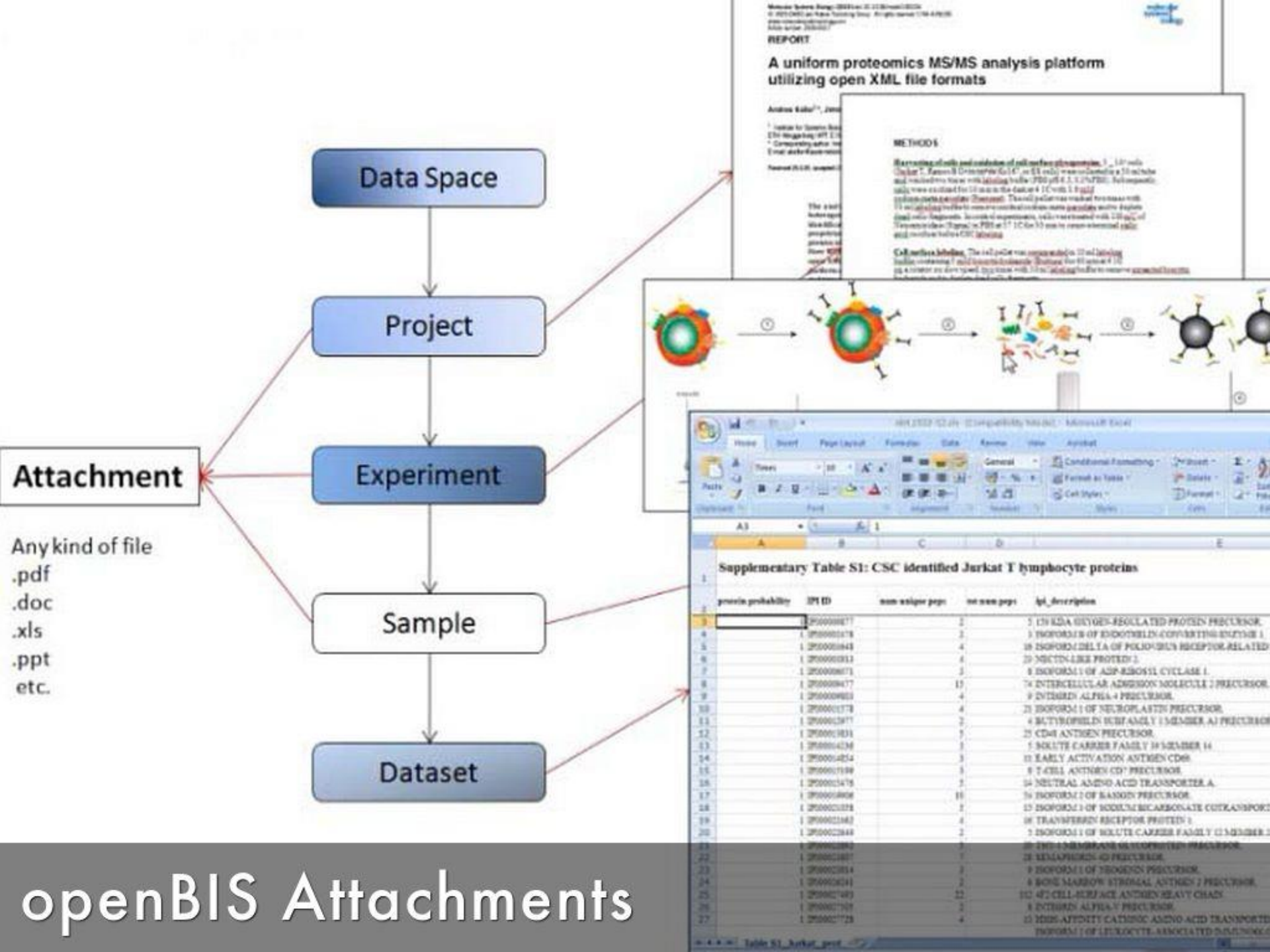


openBIS



openBIS Structure





VRE for DH

- SALSAS (@ UNIBAS):
System for Annotations
& Linkage of Sources in
Arts & Humanities



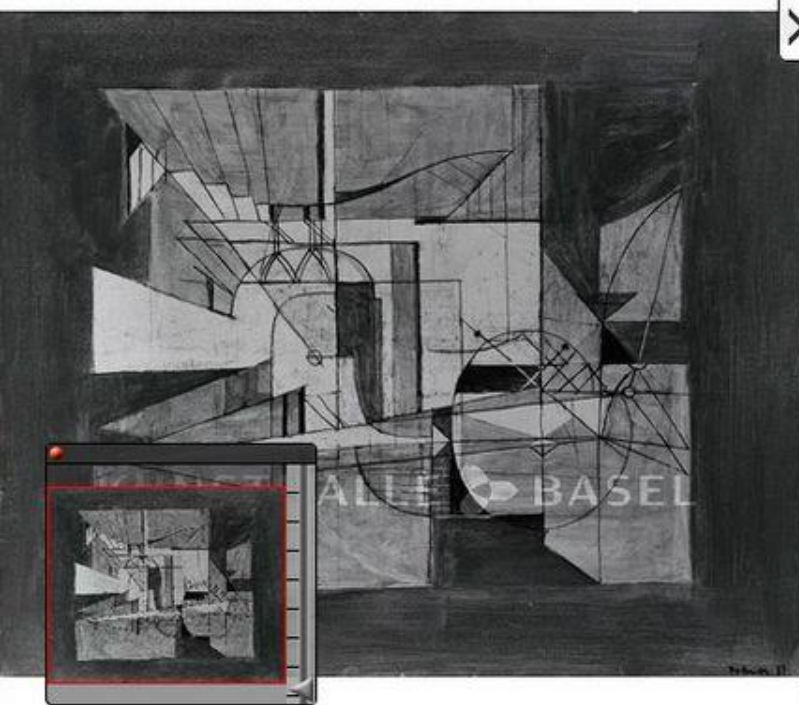
GRAPH

searchresult

Total of 14 hits

1964_03_003

818x678



Photo

Photo:

Descriptive Metadata

Identifier :

1964_03_003

Creator :

Bodmer, Walter

Contributor :

Kunsthalle Basel

Ausstellungstitel :

Walter Bodmer / Otto Tschumi / Teruko Yokoi

Ansicht :

Einzelwerk

Technik :

Schwarz-Weiss Silberbromid Abzug

Fotograf :

Moeschlin + Baur

searchresult

Total of 425 hits

_info _type

Book

Value

Title : [Das N

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Supplement

Title : UE Ver

Person

Name : Egloff

Person / Institution

Name : Egloff

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

Identifier : 19

Photo

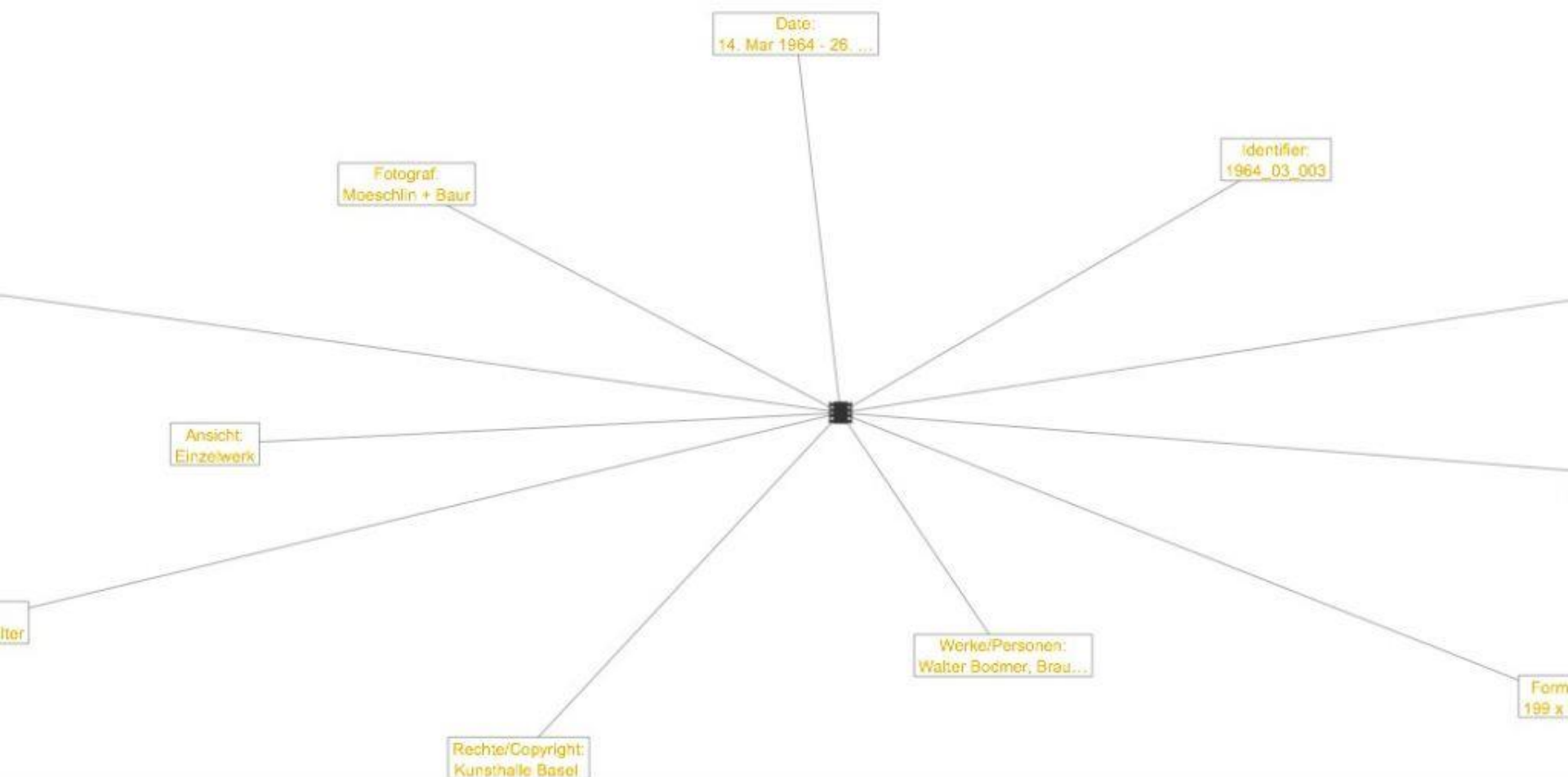
Identifier : 19

über Odyssee und Argonautika 1920, Habilitation 1926, 1933 ab. und 1942 o. Prof. für klass.

Altertumswissenschaft mit bes. Berücksichtigung der antiken Volkskunde. Mitglied div. Gelehrter

Gesellschaften, Mitglied der Akademien Heidelberg und Oslo

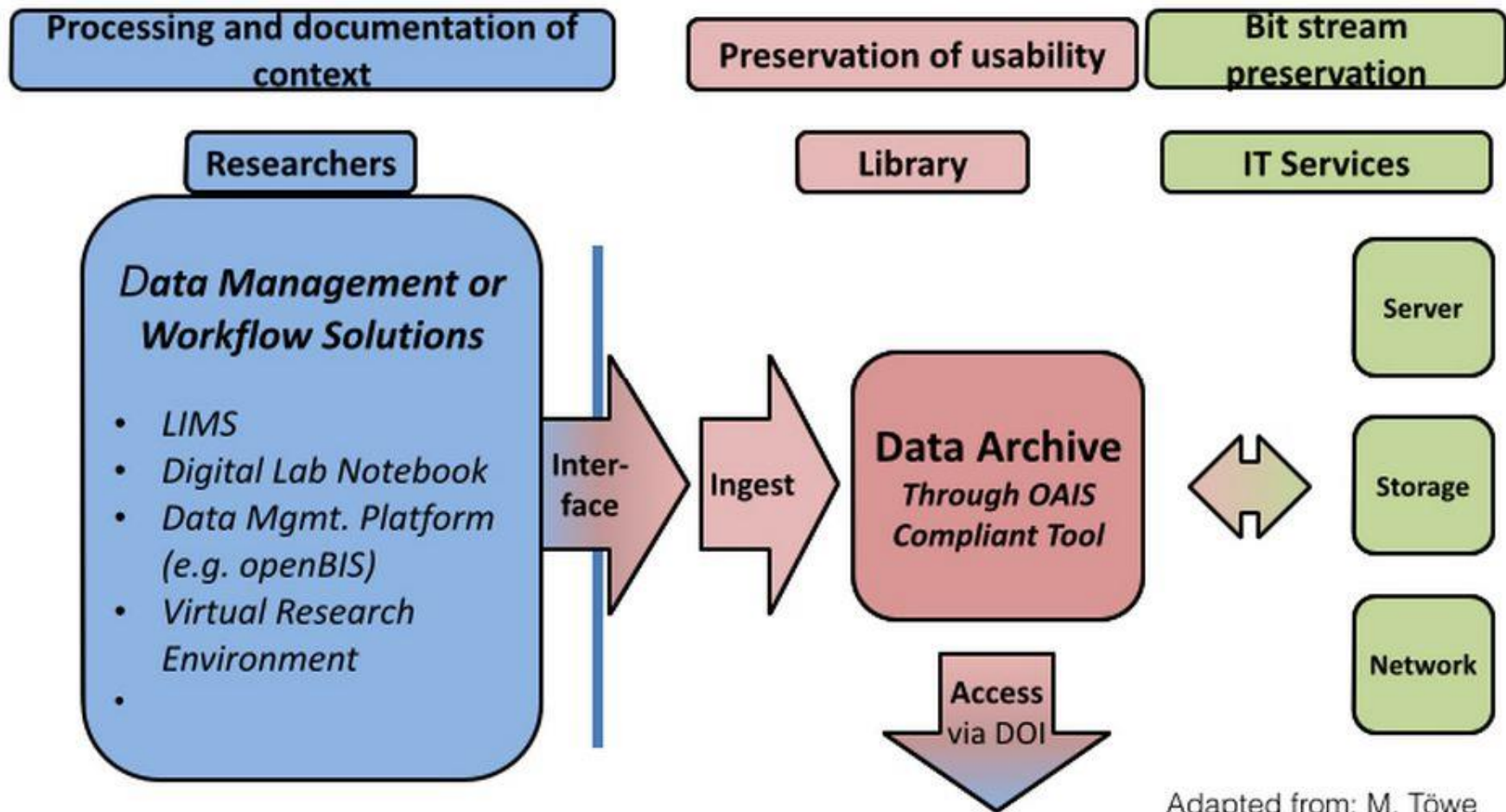
searchresult Supplement searchresult Person / Institution 1964_03_003



Salsah RDF Structure

2. Data Preservation & Sharing

- Eprints
- Rosetta
- Fedora Commons
- Docuteam Packer
- Invenio



Adapted from: M. Töwe
(oct. 14)

General Workflows

Processing and documentation of context

Researchers

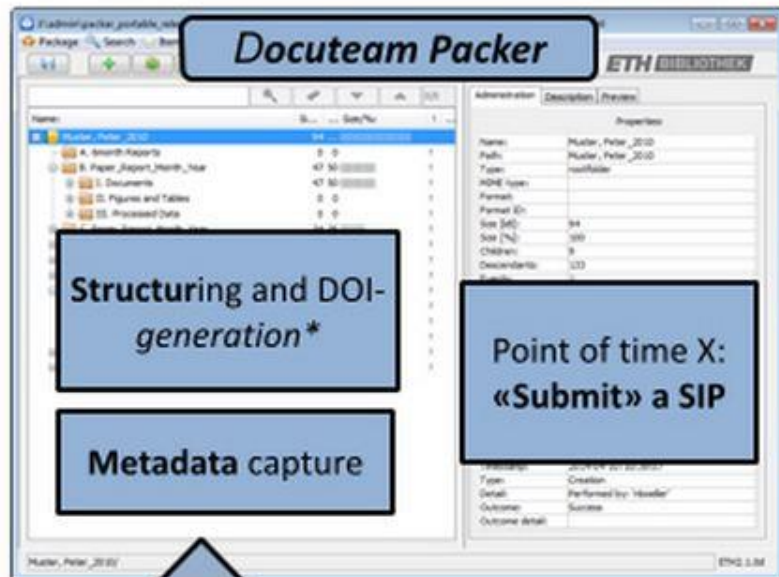
Preservation of usability

Library

Bit stream preservation

IT-Services

Docuteam Packer



Structuring and DOI-generation*

Metadata capture

Point of time X:
«Submit» a SIP

Filesystem
(networked, local)
Selection for
archiving

Submission
docuteam feeder

ETH Data Archive

Delivery
via DOI

Access via
Knowledge Portal

Server

Storage

Network

Adapted from: M. Töwe
(oct. 14)

"Small-Data" Workflow

Challenging Issues

- Genericity vs. specificity of the tools
- UI: How to intuitively handle complex datasets?
- Business plans

Directions:

1. Place kit on FIRM surface.
2. Press firmly down on circle of kit.
3. Repeat step 2 as necessary, or until unconscious.
4. If unconscious, cease stress reduction activities.

DLCM Contacts

- pierre-yves.burghi@unige.ch
(project manager)
- eliane.blumer@unige.ch
(project coordinator)