

*“The question is, how are we to make a leap from tangible things that themselves have no semantic or other meaning whatsoever, to the intersubjective sense of **collectivity** that they purported to embody?”*

*The whole endeavour seems suspicious, and not least because it is clear that ‘**connectivity**’ does not automatically entail ‘**collectivity**’, even if **collectivity** does require a minimum threshold of **connectivity**. But it would be a serious error, as Emile Durkheim writes; “if the moral concentration of a community were always judged according to the degree of physical concentration that it represented. Roads, railways, etc can serve commercial exchanges better than they can serve the fusion of populations, of which they can give only a very imperfect indication” (Durkheim 1982).”*

Contours of External and Internal in EU-Russia Relations.

Cross-biobank data integration and harmonisation

Creating shared value

Maria Krestyaninova, PhD
Uniquer Sarl, Switzerland
Institute for Molecular Medicine Finland

Content

- Looking for an interoperability solution: knowns and unknowns
- Design of various harmonisation and integration platforms
- Examples

Trend in contemporary research innovation: cross-disciplinary collaboration

- locating most informative samples
and data sets
- powering up large scale studies

IT services and communication impacting collaborative research:

- Speed up discovery
- Formation and sustaining a partnership

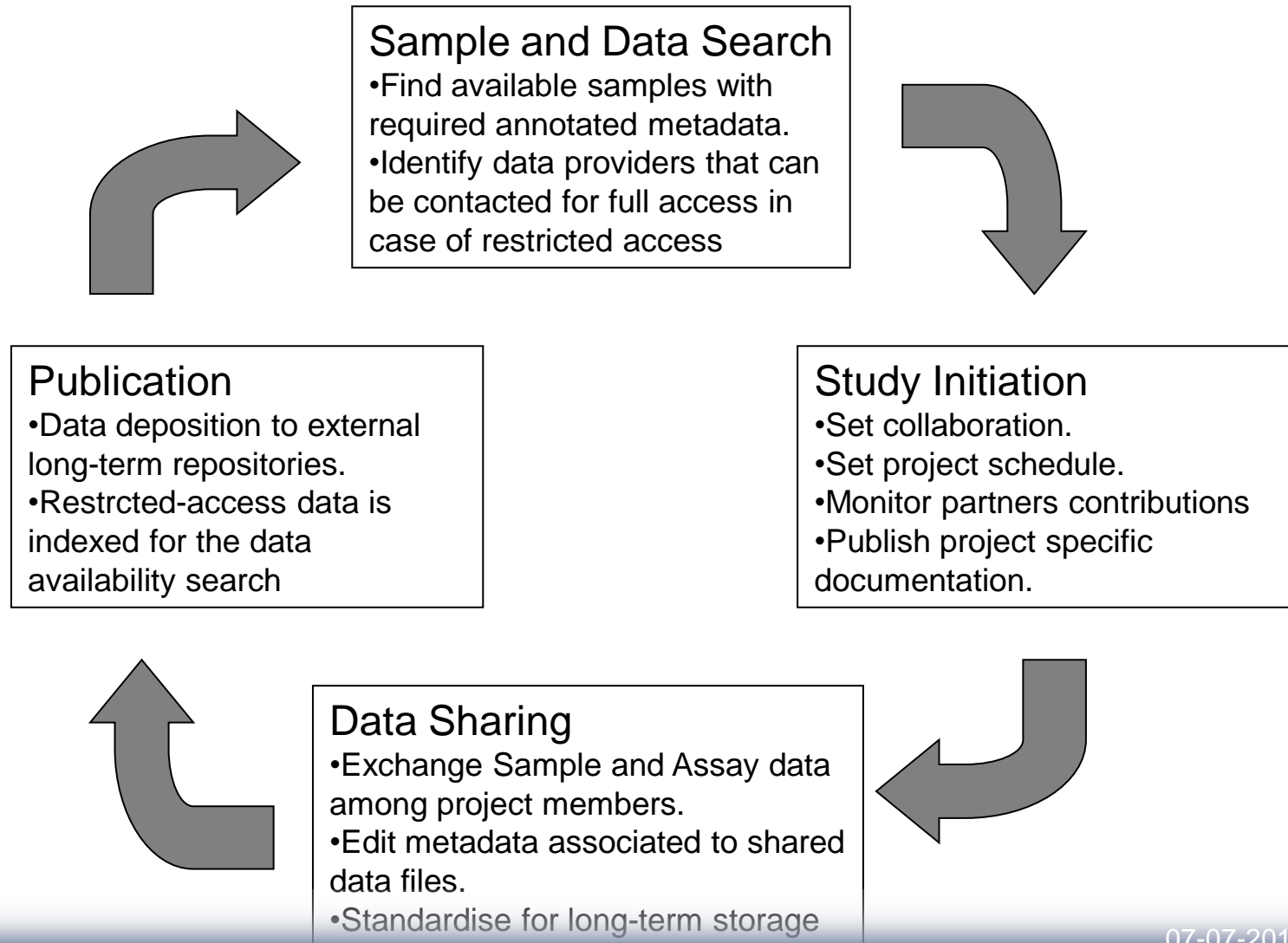
Biobank interoperability

- Value creation for large sample collections is conditional to cost-efficient communication infrastructure
- Cross-organisational data exchange and harmonisation is not a problem specific to biobanks and biorepositories
- Accurate representation of the ***sample collection content*** and of a ***researcher's intention*** (semantically, visually and ethically) is the key to formation of and sustaining a partnership

State of the art:

- Use-case
- Standards
- Two well-defined problems: semantic and data access
- IT solutions

Research flow

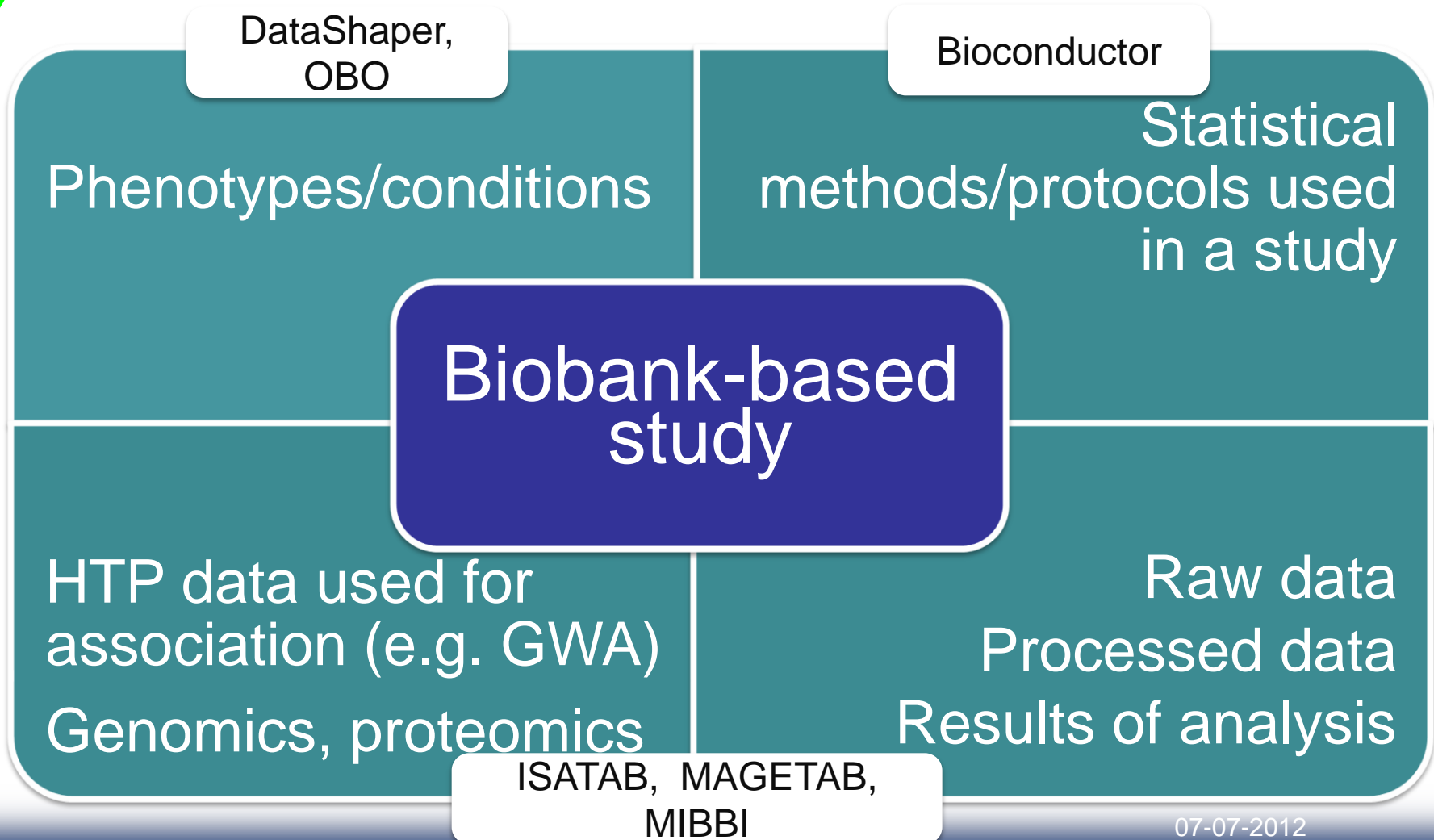


Use-cases

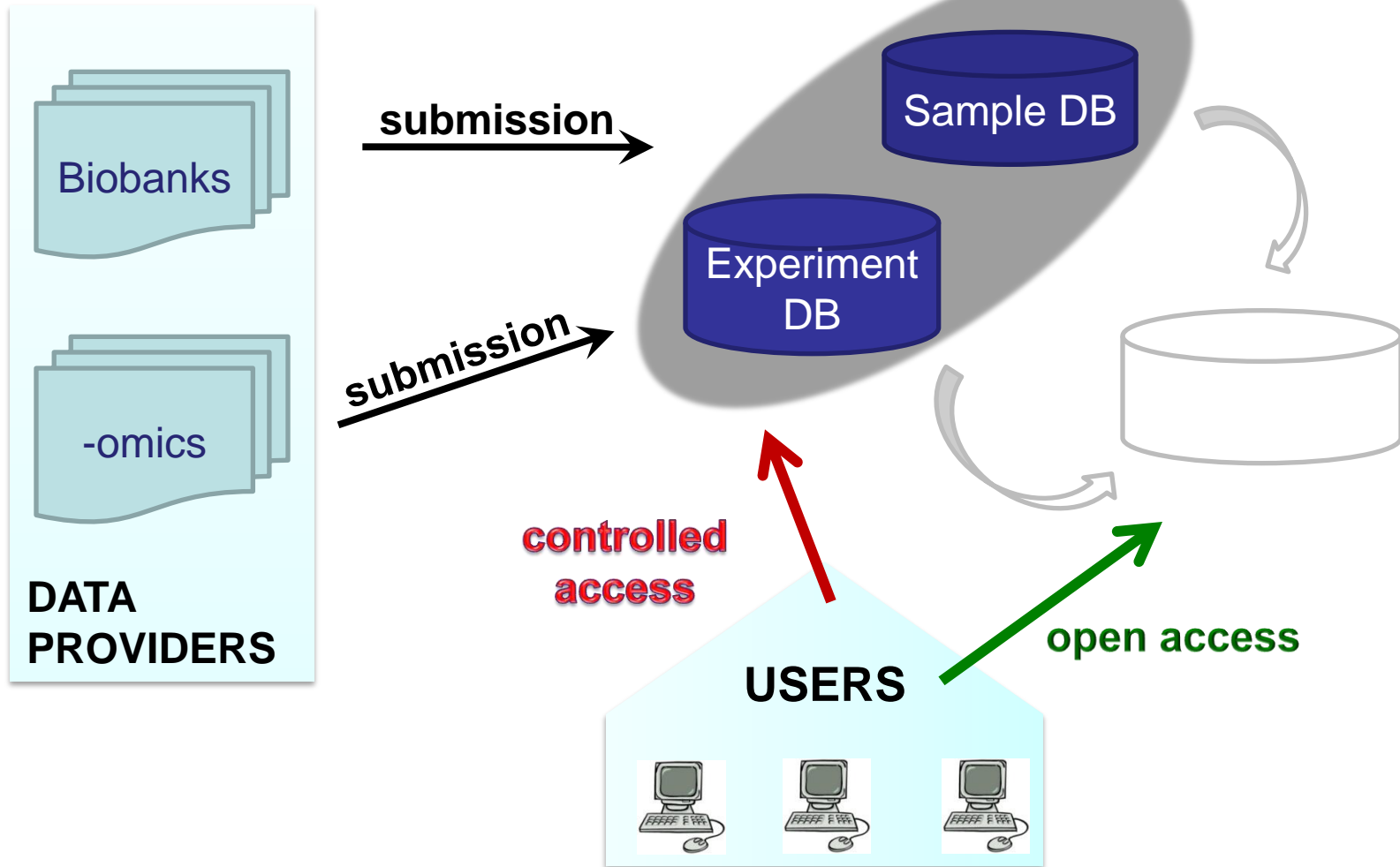


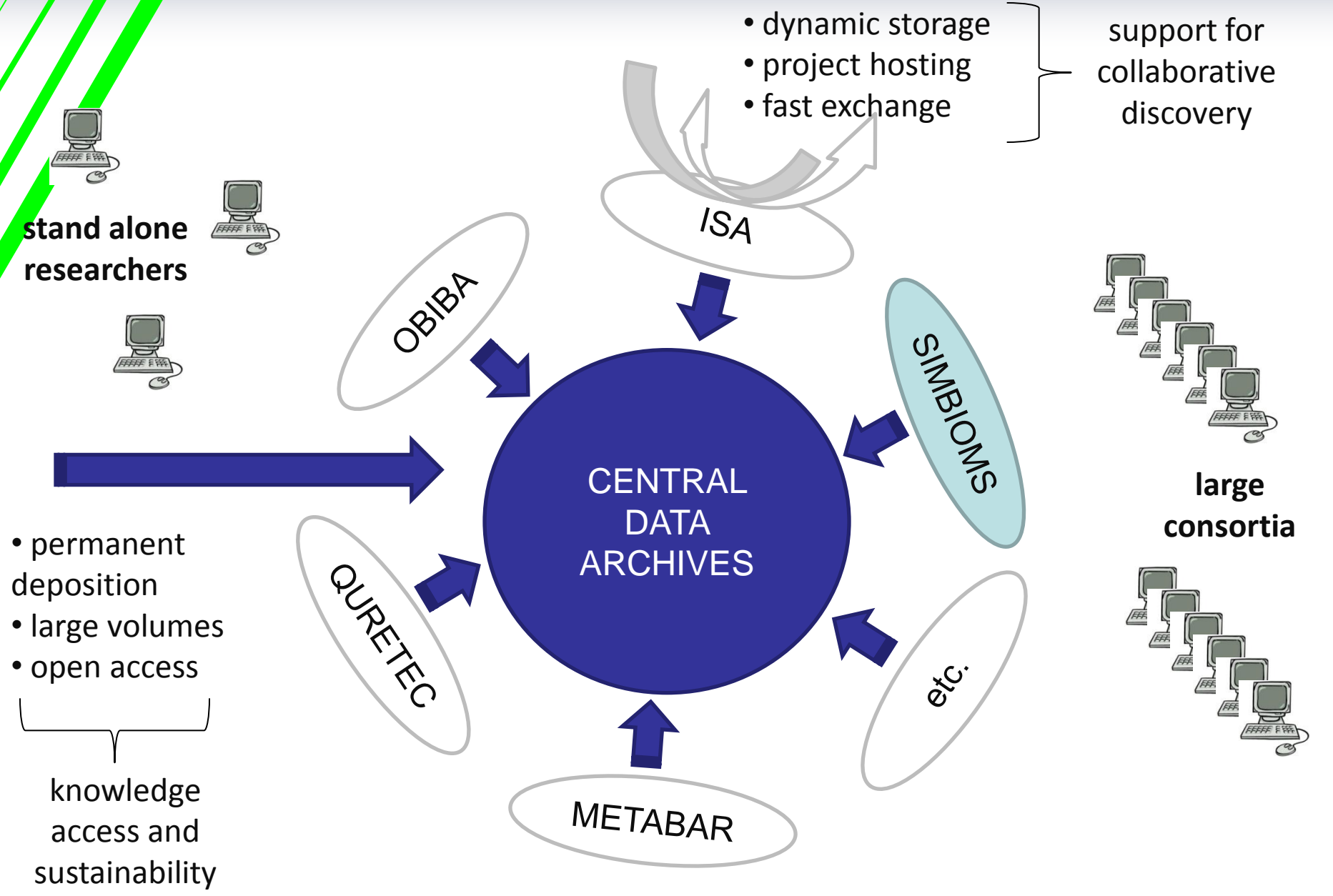
- Sample and assay tracking, house keeping data management
Internal biomaterial management: collection, distribution and QA
Users: local biobank maintenance and distribution team
Example: OBIBA, any LIMS solution
Data analysis, internal
- Data access for analysts: advanced search, analytical and visualization tools
Users: statisticians and epidemiologists
Example: Bioconductor
Harmonisation, annotation and indexing of data
- Data enrichment and annotation; mapping to standard ontologies and vocabularies
Users: data managers and epidemiologists world-wide
Example: DataSHaPER, SAIL, etc
- Controlled data and/or metadata release for collaboration enhancement
Selective and secure data sharing with collaborators and scientific community
Users: potential and existing collaborators, external
Example: EGA, SAIL, SIMS/AIMS, ISATAB

Reporting requirements and standards



System overview





State of the art:

- Use-case
- Standards
- Two well-defined problems: semantic and data access
- IT solutions

Anything still missing?

HOW TO CHOOSE AN INTEROPERABILITY SOLUTION?

Researcher's view of interoperability

- most informative samples
- sample size
- study design

IT architect perspective:

- metadata
- semantic enrichment
- standards

New angle...

- Who are the interacting parties?
- What are the communications?
- What kind of support is required?

Who has looked into the communication trends in collaborative research?

Media Engineering

Open Design

Open Innovation

Organisational research

John Christopher Jones (John Chris Jones) is a Welsh [designer](#). He was born in 1927, in [Aberystwyth, Wales](#). He studied [engineering](#) at the [University of Cambridge](#), and went on to work for [AEI](#) in [Manchester, England](#). His 1970 book *Design Methods* is considered a major text in [design](#).

Collaborating with engineers, Jones advocated [ergonomics](#) and the consideration of user-centred issues not part of [engineering](#) skills and attitudes at the time. When the results of his ergonomic studies of user behavior were not utilized by the firm's [designers](#), Jones set about studying the [design process](#) being used. Jones was also frustrated with the superficiality of [industrial design](#) at the time and become involved with [ergonomics](#).

[Design methods](#) as an area was driven by:

- Inability to balance individual, group, societal, and ecological needs;
- Lack of purpose, order, and human scale;
- Aesthetic and functional failure in adapting to local physical and social environments;
- Development of materials and standardized components that were ill suited for use in any specific application;
- Creation of artifacts that people did not like

Jones wasn't actually addressing [design](#) as presently conceived. He set out an entirely original philosophy of design—one that questioned the aims, goals and purposes of designing. He stated that one of the reasons why he focused on Design Methods was

"... it's not another way of doing design, you see, it's a way of doing what designers don't do at all."

At the end of the 1950s he published an article "A Systematic Design Method" articulating ways to integrate ergonomic data into the engineering design process. His emerging ideas about Design Methods was to integrate both [rationality](#) and [intuition](#)—a common thread in the formalization of Design Methods and how it was interpreted by other groups.

He also realized that designers needed to move out of focusing on [expression](#)^{[\[disambiguation needed\]](#)} and modes of production and begin to address the definition of a problem to be solved. He commented that

"the future job of a designer is to give substance to new ideas while taking away the physical and organizational foundations of old ones. In this situation, it is nonsense to think of designing as the satisfaction of existing requirements. New needs grow and old needs decay . . ."

Books

[\[edit\]](#)

- Jones, John Christopher, *Design Methods: seeds of human futures*, John Wiley & Sons Ltd., London, 1970; 2nd edition, John Wiley & Sons Ltd., 1992
- Jones, John Christopher, *Designing Designing* (London: Architecture Design and Technology Press), 1991

Design Methods is a broad area that focuses on:

[Divergence](#) – Exploring possibilities and constraints of inherited situations by applying [critical thinking](#) through qualitative and quantitative research methods to create new understanding (problem space) toward better design solutions

Transformation – Redefining specifications of design solutions which can lead to better guidelines for traditional and contemporary design activities (architecture, graphic, industrial, information, interaction, et al.) and/or multidisciplinary response

[Convergence](#) – [Prototyping](#) possible scenarios for better design solutions that incrementally or significantly improve the originally inherited situation

[Sustainability](#) – Managing the process of exploring, redefining and prototyping of design solutions continually over time

[Articulation](#) - the visual relationship between the parts and the whole.

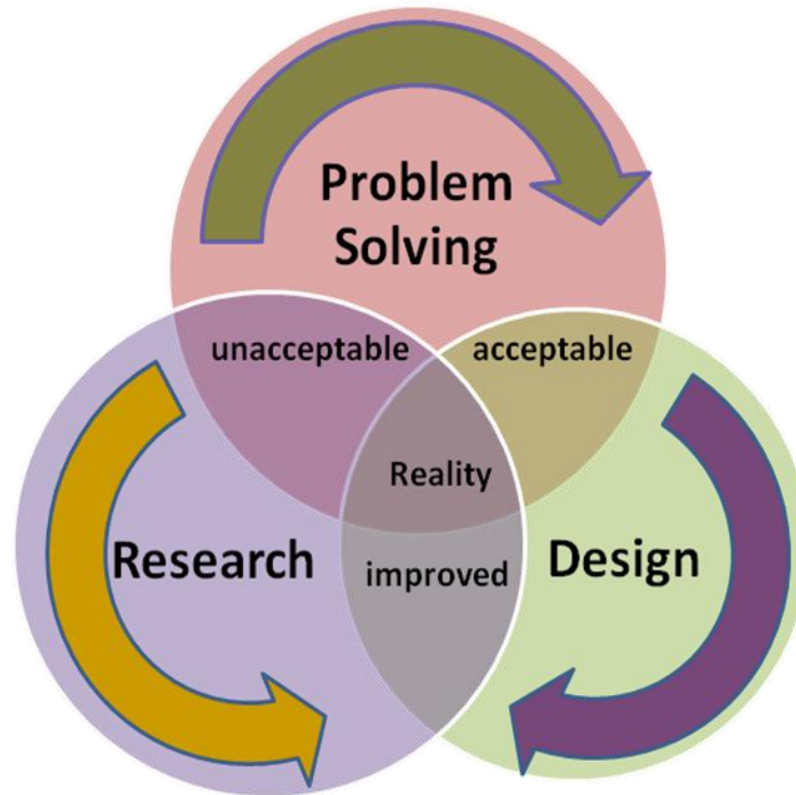
The goal of design methods is to gain key insights or unique essential truths resulting in more [holistic](#) solutions in order to achieve better experiences for users with products, services, environments and systems they rely upon. Insight, in this case, is clear and deep investigation of a situation through design methods, thereby grasping the inner nature of things intuitively.



John Chris Jones



UNDERSTANDING REALIZATION ACTIVITIES



Research: aiming at understand reality

Problem solving: aiming at utilize reality

Design: aiming at improving or altering reality

Metadesign

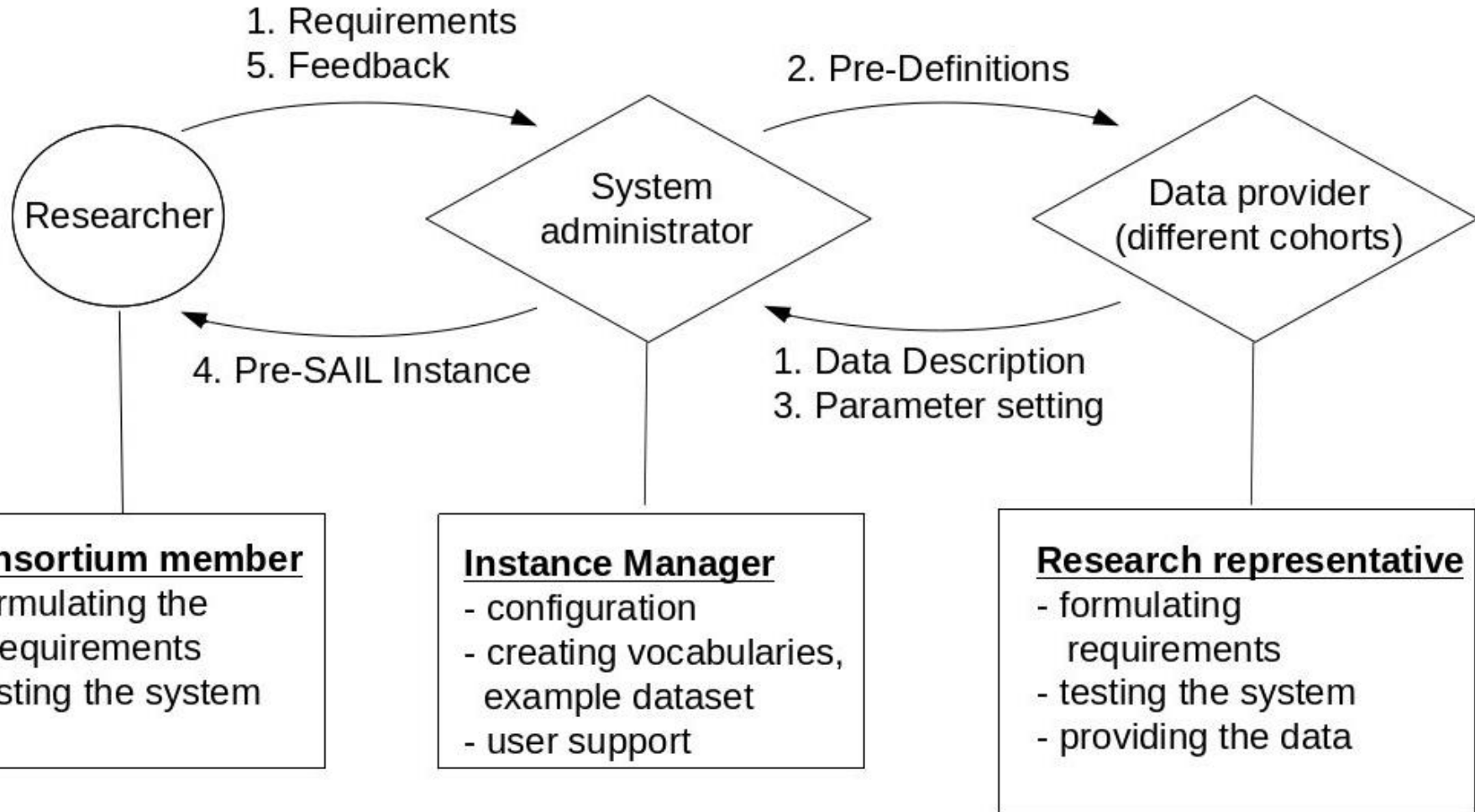
Participation level (individual/ collective)	Analysis	Concept design	Concept communication	Distribution	End-of-life
none					
indirect					
consultative					
Shared control					
Full control					

Designing infrastructure for biobanks

- Who are the players?
- What are the interactions?
- What kind of communication enhancement could Information Technology provide to biobanks

Let's zoom in...

Communication flow



by Jörn Dietrich, June 2011, Berlin

And zoom in...

Problematics

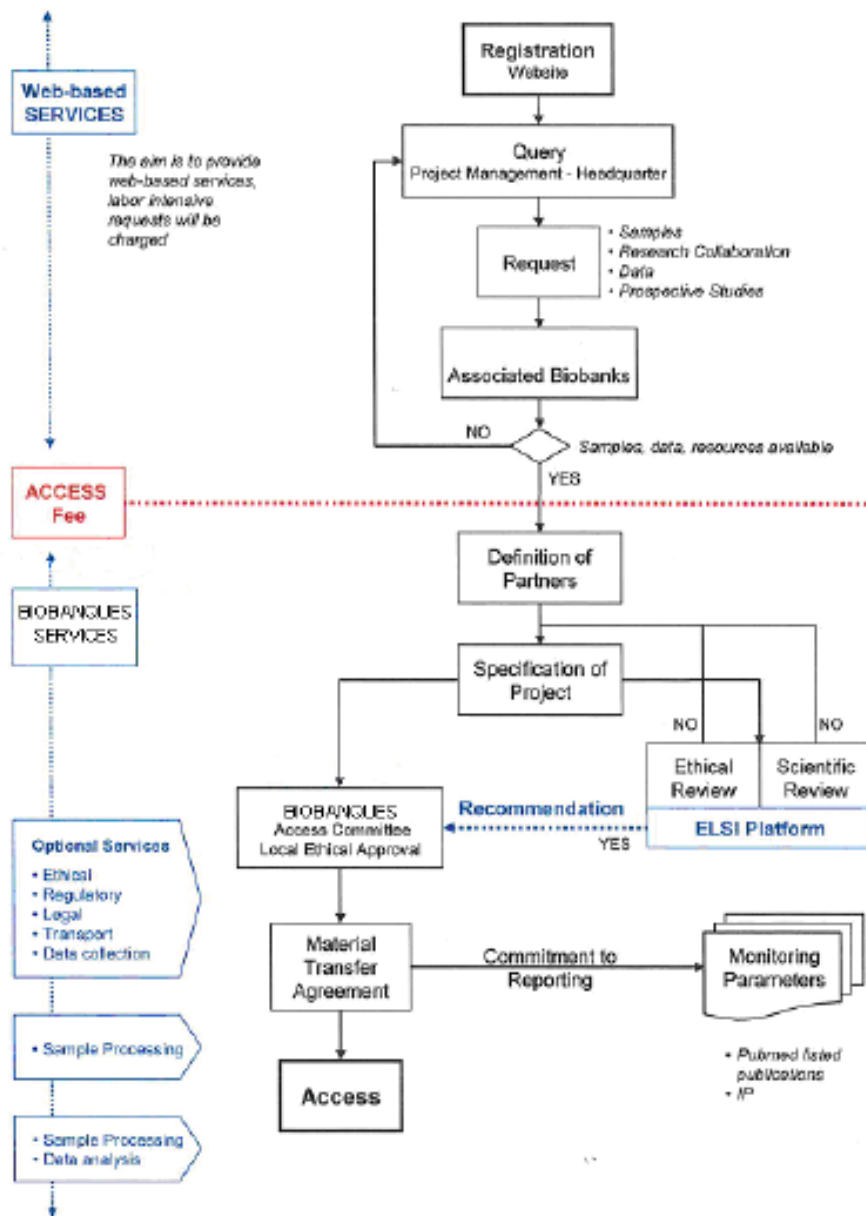


High diversity of data sources, formats, description and of research question and methodologies.

Complex requirements for data access

Modelling communication flow

- Facilitate formation of collaboration
- Speed up design of studies and discovery
- Support data re-annotation, curation and mapping
- Contribute to data Q/A, consistency and sanity checks



PART III

EXAMPES

International initiatives aiming to enhance biobank interconnectivity

- Public Population Project in Genomics (P3G)

- www.p3g.org

- Biobanking and Biomolecular Resources Research Infrastructure (BBMRI)

- www.bbmri.eu

[VIEW REQUESTS](#)

SEE SPECIMEN REQUESTS

[POST REQUESTS](#)

MAKE SPECIMEN REQUESTS

[BIOBANK DIRECTORY](#)

GLOBAL LISTING OF BIOBANKS

[SUPPORT](#)

BIOBANKS

GLOBAL DIRECTORY OF BIOBANKS, TISSUE BANKS AND BIOREPOSITORIES

Following is the world's most comprehensive directory of biobanks, tissue banks and biorepositories that collect blood, whole blood, buccal swab, DNA, RNA, protein, cell lines, marrow, plasma, serum, RBC, white cells, buffy coat, fluid, urine, stem cells, and solid tissue such as tumor, tumour and biopsy materials spanning all types of common and rare pathologies and indications including Alzheimer's, basal cell carcinoma, bladder cancer, bone cancer, brain cancer, breast cancer, cerebrospinal fluid, amniotic fluid, colorectal cancer, colon cancer, hodgkins and non-hodgkins lymphoma, kidney/renal cancer, leukemia, multiple sclerosis, liver cancer, lung cancer, melanoma, myeloma neuroblastoma, neurodegenerative diseases, ovarian cancer, pancreatic cancer, prostate cancer, synovial fluid, urinary cancer and more.

Email us at specimencentral@gmail.com to request inclusion of your biobank in SpecimenCentral.

Global Regions

[EUROPEAN BIOBANKS](#)[NORTH AMERICAN BIOBANKS](#)[ASIAN BIOBANKS](#)[AUSTRALIAN BIOBANKS](#)[MIDDLE EAST BIOBANKS](#)[ANIMAL & PLANT BIOBANKS](#)

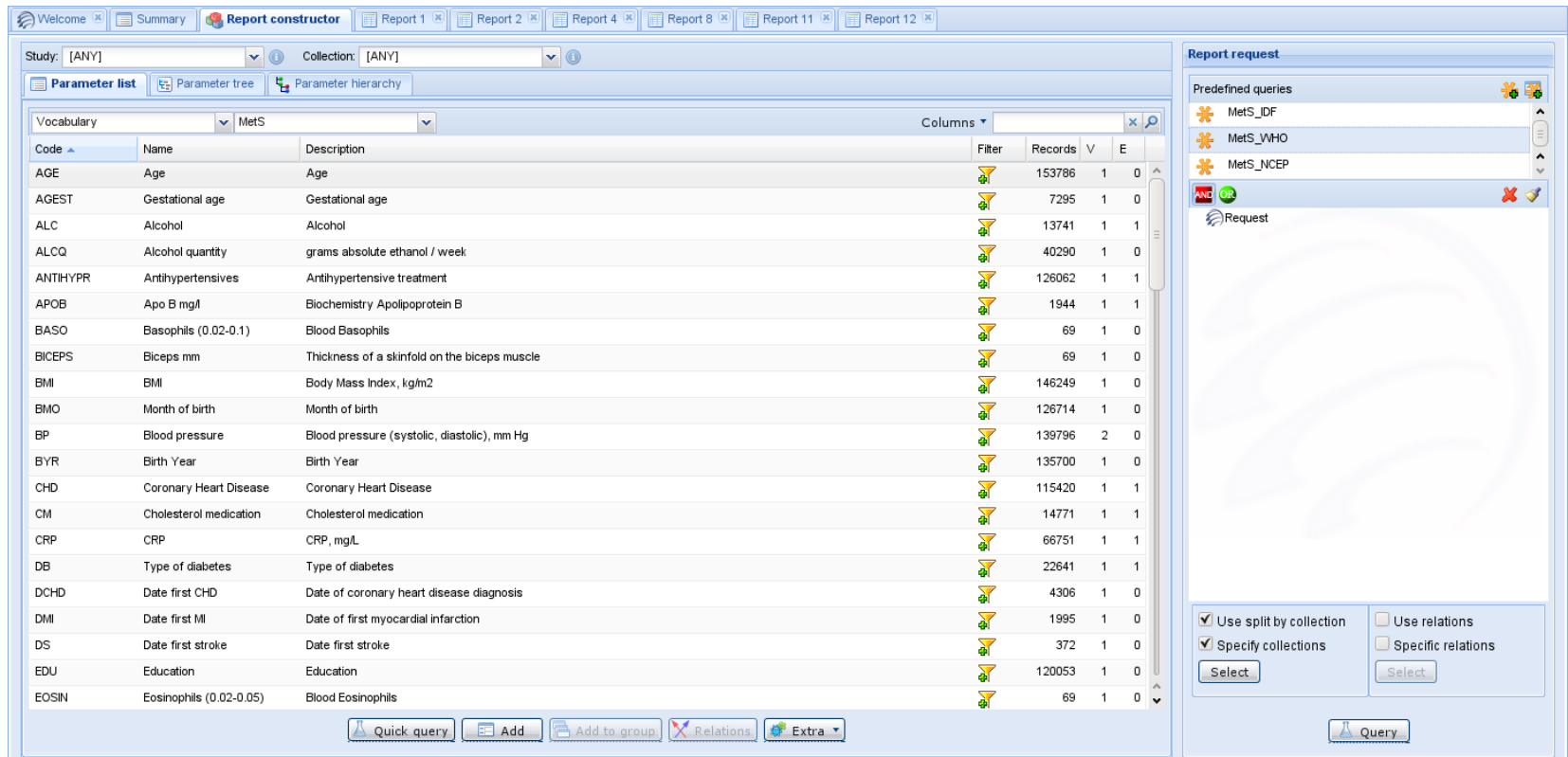
[Contact SpecimenCentral](#)

For immediate requests, please contact us via comments, email, or social media.
[specimencentral@gmail.com](#)

(WE INVITE YOU TO IMPROVE THE DIRECTORY)

Biobank metaportal: in search for most informative samples

Web-based metaportal



Study: [ANY] Collection: [ANY]

Parameter list | Parameter tree | Parameter hierarchy

Vocabulary: MetS

Code	Name	Description	Filter	Records	V	E
AGE	Age	Age		153786	1	0
AGEST	Gestational age	Gestational age		7295	1	0
ALC	Alcohol	Alcohol		13741	1	1
ALCQ	Alcohol quantity	grams absolute ethanol / week		40290	1	0
ANTHYPR	Antihypertensives	Antihypertensive treatment		126062	1	1
APOB	Apo B mg/l	Biochemistry Apolipoprotein B		1944	1	1
BASO	Basophils (0.02-0.1)	Blood Basophils		69	1	0
BICEPS	Biceps mm	Thickness of a skinfold on the biceps muscle		69	1	0
BMI	BMI	Body Mass Index, kg/m2		146249	1	0
BMO	Month of birth	Month of birth		126714	1	0
BP	Blood pressure	Blood pressure (systolic, diastolic), mm Hg		139796	2	0
BYR	Birth Year	Birth Year		135700	1	0
CHD	Coronary Heart Disease	Coronary Heart Disease		115420	1	1
CM	Cholesterol medication	Cholesterol medication		14771	1	1
CRP	CRP	CRP, mg/L		66751	1	1
DB	Type of diabetes	Type of diabetes		22641	1	1
DCHD	Date first CHD	Date of coronary heart disease diagnosis		4306	1	0
DMI	Date first MI	Date of first myocardial infarction		1995	1	0
DS	Date first stroke	Date first stroke		372	1	0
EDU	Education	Education		120053	1	0
EOSIN	Eosinophils (0.02-0.05)	Blood Eosinophils		69	1	0

Report request

Predefined queries

- MetS_IDF
- MetS_WHO
- MetS_NCEP

Request

Use split by collection Use relations

Specify collections Specific relations

Quick query | Add | Add to group | Relations | Extra

Query

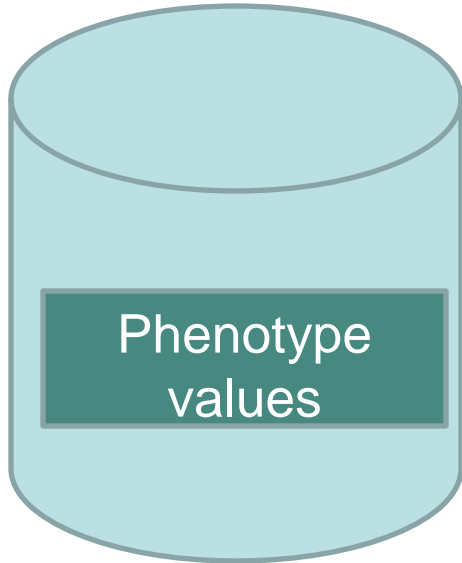
sail.simbioms.org:

~200 000 samples
from 14 biobanks

Gostev et al, Bioinf, 2010

available for researchers to construct metastudies

SIMS – Sample Information Management System (access control)

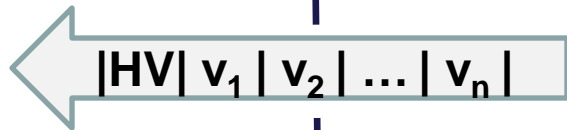


Sample matrix

HV	v ₁	...	v _n
s ₁	0	...	1
s ₂	1	...	
...
s _k	male	...	30

PHASE II

availability data or real values



only HV



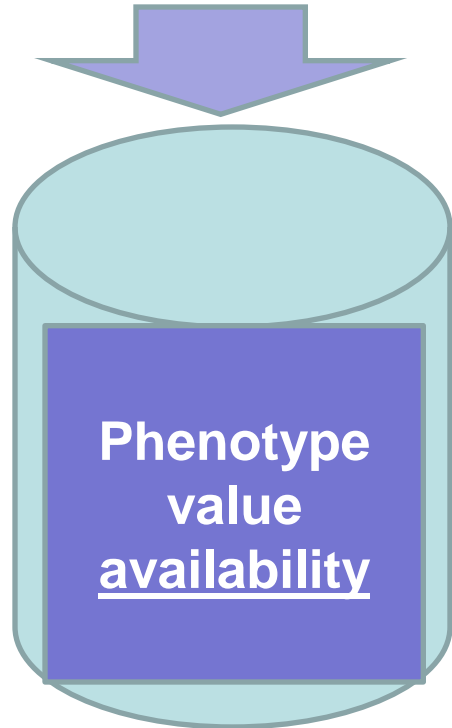
only availability data

PHASE I

HV – harmonized vocabulary
OV – original vocabulary

Variable matrix

HV	OV ₁	...	OV _m
v ₁	Mappings between HV and OV		
v ₂			
...			
v _n			



SAIL – Sample Availability Information System (public). Search and sample counts.



- Collections summary
- Report constructor
- Tutorial
- Source

Central indexing of data stored locally

- Heavy or protected data remains in original location
- Availability of data points or samples is indexed centrally against Variables of Interest
- Amount and location of data relevant to a research question is estimated online

<http://sail.simbioms.org>

Currently

- 3 production instances
- ~200 000 samples characterised with more than 100 variables of interest
- >20 data contributors to the common data index of biomaterials across-europe
- collaboration with biospecimen catalogues and ontology initiatives

Next

- tighter integration with suppliers of standard ontologies and vocabularies
- building on longitudinal data handling expertise in other domains
- wider data contributor circle through interaction with national biobanking initiatives

Media for collaborative research and innovation

- semantically-enhanced inter-disciplinary interface
- Biobanks adequately equipped for fine-tuning of data visibility and for the data leakage prevention
- means to discover and request data

technical

- recognition of diversity and various roles
- awareness of collaborative ecosystem
- long-term vision (sustainability, legacy)

organisational

After-word



In order to cover currently existing gap between

- *IT infrastructure technical capabilities [**connectivity**]*
and
- *the internal and external communication needs of research collaborative ecosystem [**collectivity**]*

we shall bring together IT architects, designers, data managers, communication engineers and social scientists.



Institute of Mathematics and Computer Science
University of Latvia

- Juris Viksna
- Martins Opmanis
- Andris Zarins



Institute for Molecular Medicine Finland
Nordic EMBL Partnership for Molecular Medicine

- Teemu Perheentupa
- Jani Heikkinen
- Timo Miettinen



Uniquer

- Joern Dietrich
- Russell Vincent
- Yulia Tammisto

EMBL-EBI



- Ugis Sarkans
- Stathis Kanterakis



**Karolinska
Institutet**

- Ola Spjuth

SIMBIOMS in collaborative biomedical research initiatives

Project	Goal/Description	Funded by	Simbioms team involvement
Strategic research collaborations			
BBMRI www.bbmri.eu	Build a network of population-based biobanks, experts, and foster collaboration between them.	EC, OECD	Prototyping of data management model, use-case design, discussions.
P3G www.p3g.org	Provide advice to industry.	Canadian Gov., memberships	Leading international Informatics Working Group; discussions.
ELIXIR www.elixir-europe.org/page.php	Create a sustainable infrastructure for the storage and distribution of information produced by bioscientists.	EC	Prototyping, reports, cooperation with organisation of medical informatics committee on behalf of EBI.
TaraOceans oceans.taraexpeditions.org	3-year long circumnavigation expedition for marine genomics and climate integrative study.	CNRS, industry, potentially EC	Preliminary design of data management solution; meetings, discussions.
Services for research collaborations			
ENGAGE www.euengage.org	Genetic and genomic research for clinical application.	EC	Design, development and maintenance of dedicated data exchange services – based on SIMBioMS.
MolPAGE www.molpage.org	Biomarkers: discovery and development of novel high-throughput methods.	EC	
MuTHER	Exploration of gene expression in multiple tissues on 1000 twins associated with aging.	Wellcome Trust	
SIROCCO www.sirocco-project.eu	Study of small RNAs as regulatory cell mechanism; therapeutical applications.	EC	
CAGEKID	Kidney cancer study.	EC	
SUMMIT	Surrogate markers for vascular Micro- & Macrovascular hard endpoints for Innovative diabetes Tools	EC	

