CHIC

CHIC-CDR

A repository for managing multi-modality clinical data and its application to *in-silico* oncology

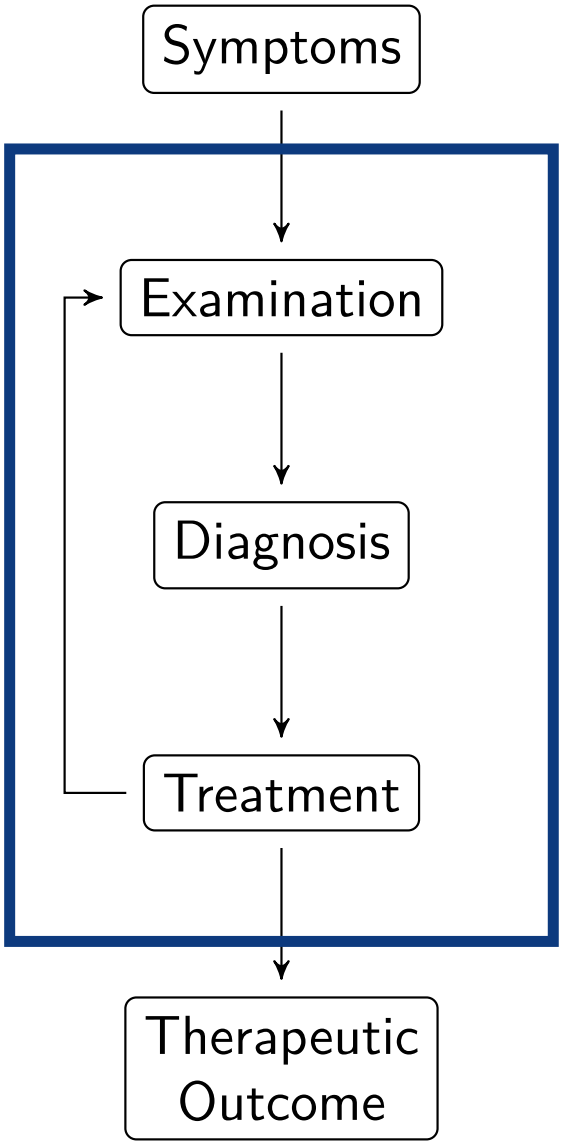
M. Kistler[1,2] R. Niklaus[1,2] D. Abler[1] P. Büchler[1]

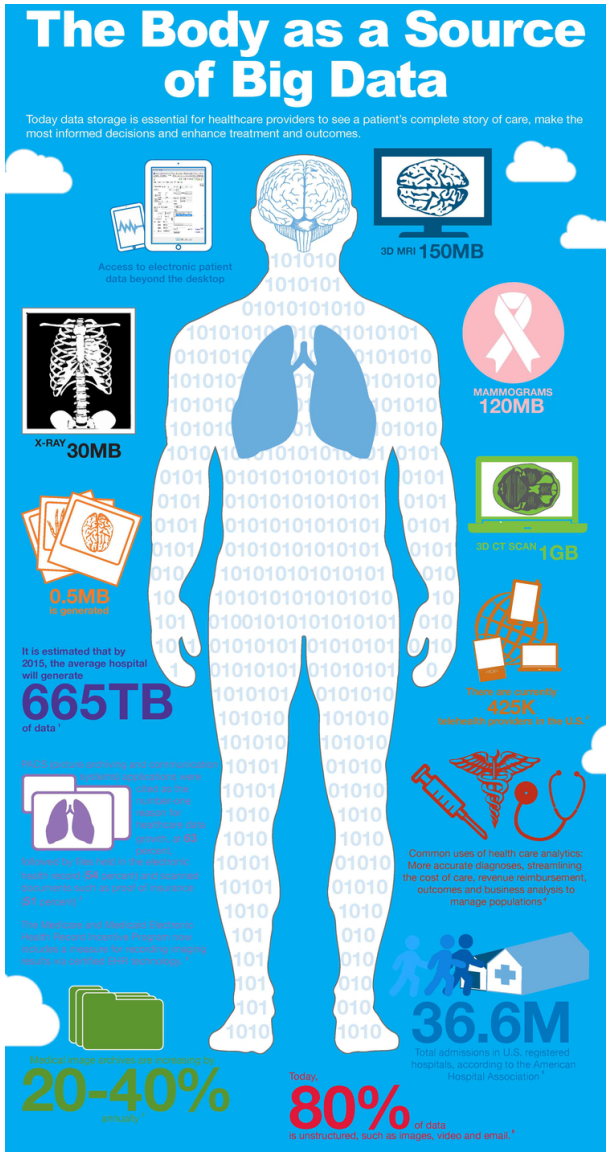
[1] Institute for Surgical Technology & Biomechanics, University of Bern, Bern, CH

[2] Swiss Institute for Computer Assisted Surgery, Courroux, CH

Contents

Clinical Data & Research





Clinical Data & Research

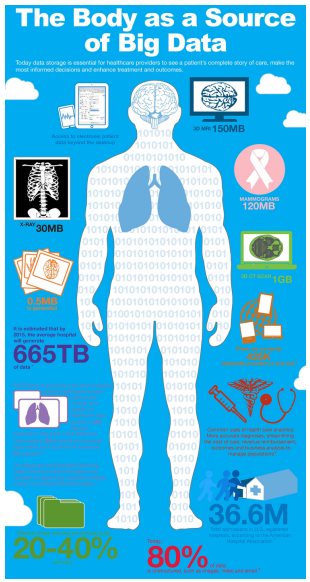
Symptoms

Examination

Diagnosis

Treatment

Therapeutic Outcome



Hypothesis

Data Analysis

Results

Symptoms

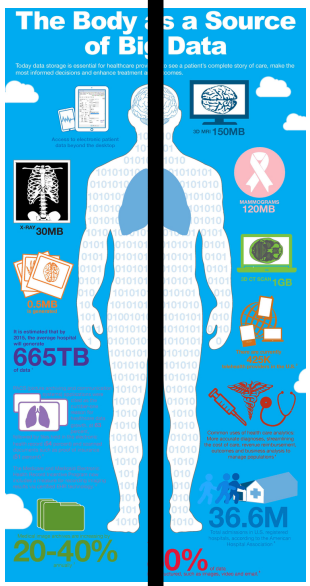
Treatment

Examination

Diagnosis

Treatment

Therapeutic Outcome



Research

Hypothesis

Data Analysis

Results

Data Sharing: Treatment vs Research

	Treatment	Research
Patient identification	yes	(pseudo-)anonymized
Access Control	+++	*scenario-specific*
Data Safety	+++	*scenario-specific*
Consent needed	n.a.	yes
Medical Device Cert.	*scenario-specific*	*scenario-specific*
Data selection	patient	patient characteristics
Search	patient history	cross-sectional & longitudinal
Importance of Quality	+++	+++
	but different criteria, scenario-dependent	

- Research comprises many different scenarios
- Data requirements differ across usage contexts
- Data sharing for treatment *and* research challenging

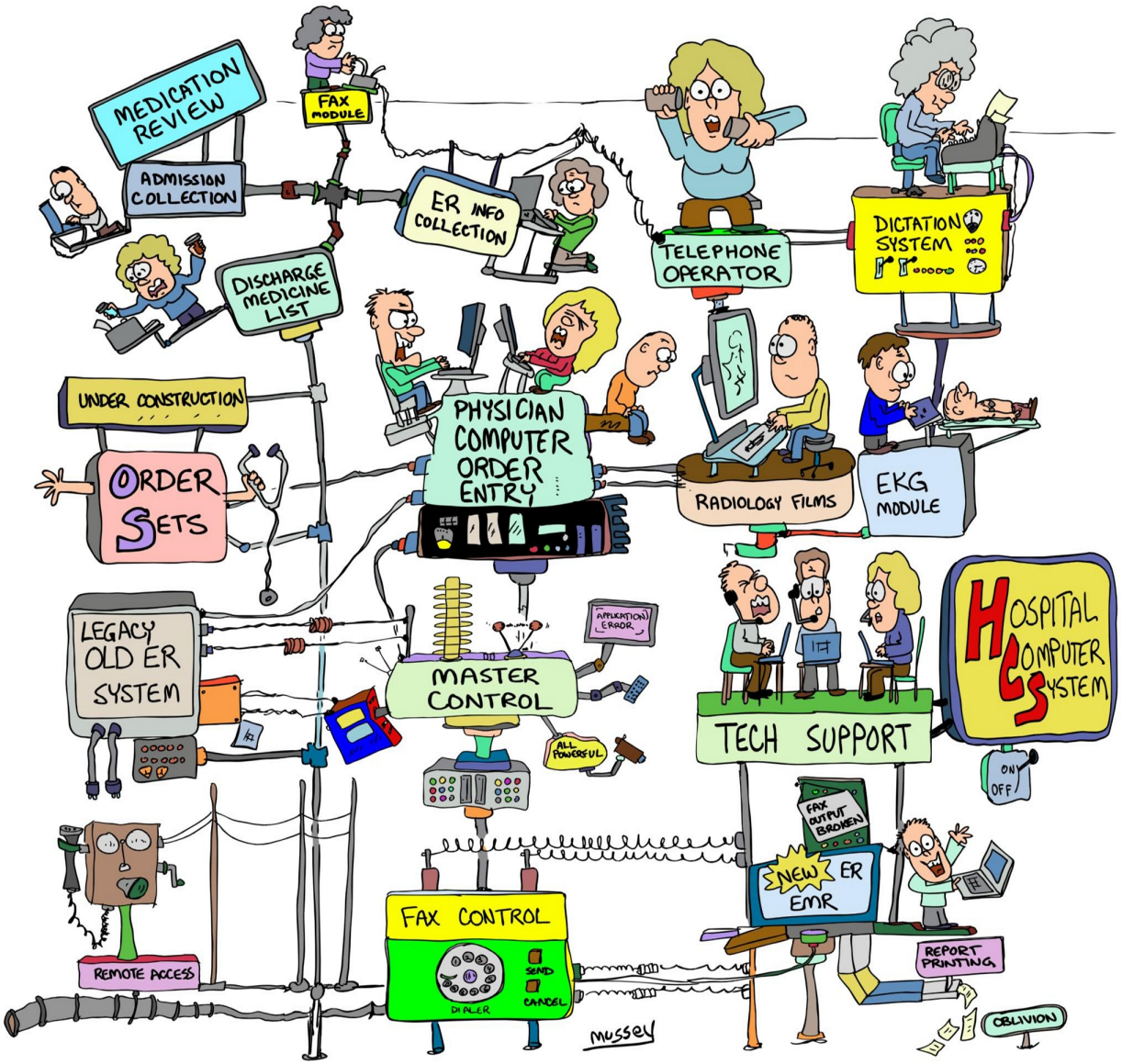
Data is prerequisite for quantititative evaluation

- Clinical Practice
  - Identification of best-practices
  - Definition of standards of care
- Scientific advancement
  - Enhanced understanding of disease progress & treatment
  - Generation of new research questions
  - Increasing reproduceability of research
- Innovation
  - Development of new diagnostic/treatment tools & methods
  - Validation & Improvement

Challenges

- Ethical & Legal
  - Data protection legislation prescribes rules for dealing with personal data
  - Different forms of "consent" handle breaches of privacy & confidentiality
- Heterogeneity of systems & formats
  - Different vendors use different standards
  - Different implementations of same standard
  - Syntactic interoperability: technical (in)compatibilities
- Diversity of practices & insufficient documentation
  - Implicit assumptions of contextual information
  - Documentation often not suited for automatic processing
  - Semantic interoperability: (in)correct interpretation of information
- Lack of mechanisms to encourage data sharing

Dedicated Data Repositories for Research?



Source: authenticmedicine.com

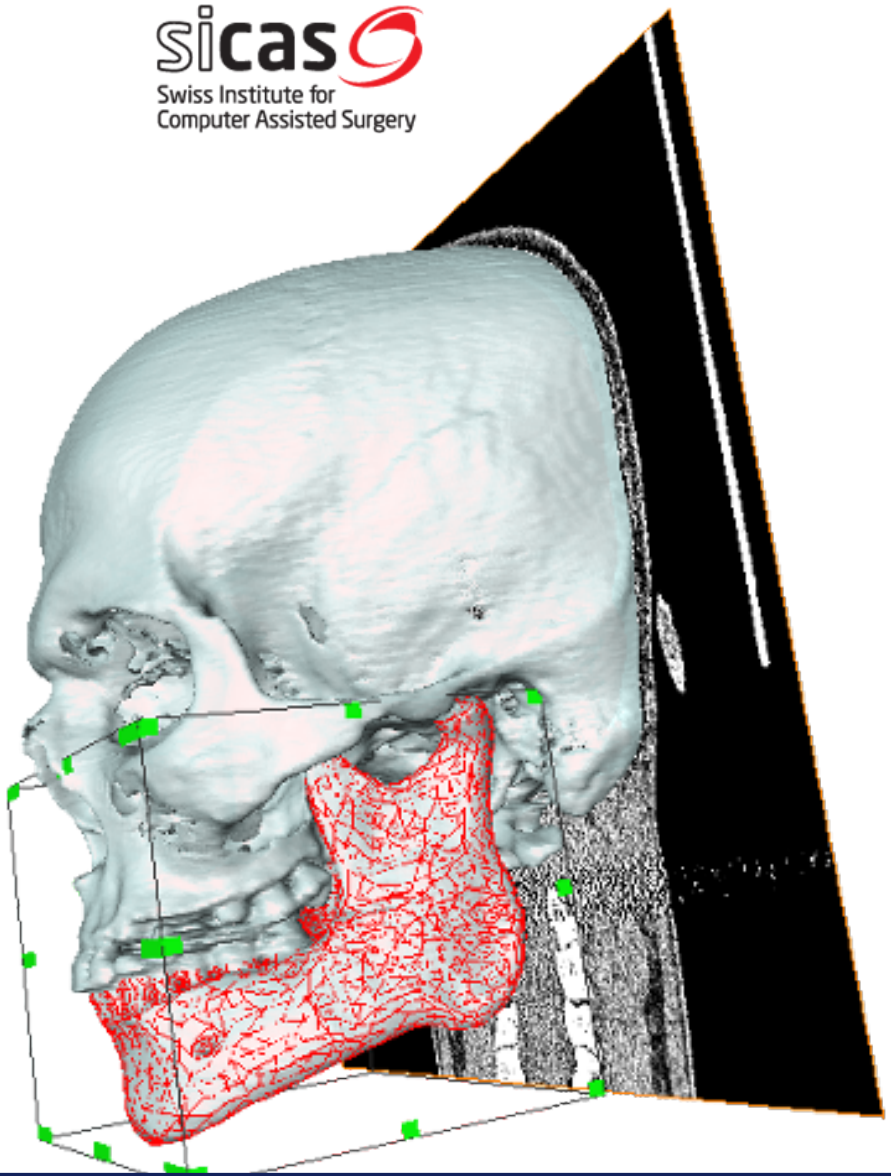
- HIS are complex
- Data spread over multiple systems
- Different primary tasks
- No integrated access

Need unique access point to relevant data.

Contents

Design Criteria

- Flexibility:
  - Custom folder-like structure for data organisation
- Accessibility:
  - Web browser & API
  - Controlled access, access levels set per data-set
  - Semantic search
- Standardisation:
  - File formats: DICOM, ITK images, CDISC, STL, Statismo
  - Public ontologies for semantic annotation



Origins & Timeline

2009–2013 Development of data repository for CO-ME project
"Virtual Skeleton Database" (VSD)
- Dicom integration & other images (ITK formats: nii, mha, etc)
- Statistical shape models
- Semantic search for anatomic structures

2013 End of CO-ME, Transfer to SICAS
"SICAS Medical Image Repository" (SMIR)

2013–now Co-development SICAS / CHIC-project
"Clinical Data Repository" (CDR)
- REST-API
- Support for (pre-)clinical data formats
- Enhanced semantic annotation & search

CO-ME
COMPUTER AIDED AND IMAGE GUIDED MEDICAL INTERVENTIONS

COmputer aided and image guided MEdical interventions

- Swiss "National Center of Competence in Research"
- Data repository for collaborative development of statistical shape models and surgical applications

Origins & Timeline

2009–2013 Development of data repository for CO-ME project
"Virtual Skeleton Database" (VSD)
- Dicom integration & other images (ITK formats: nii, mha, etc)
- Statistical shape models
- Semantic search for anatomic structures

2013 End of CO-ME, Transfer to SICAS
"SICAS Medical Image Repository" (SMIR)

2013–now Co-development SICAS / CHIC-project
"Clinical Data Repository" (CDR)
- REST-API
- Support for (pre-)clinical data formats
- Enhanced semantic annotation & search

sicas

Swiss Institute for Computer Assisted Surgery

- Support platform for collaboration in translational R&D, education, technology transfer and innovation.
- Medical interventions, therapies and health care devices.

Origins & Timeline

2009–2013 Development of data repository for CO-ME project
"Virtual Skeleton Database" (VSD)

- Dicom integration & other images (ITK formats: nii, mha, etc)
- Statistical shape models
- Semantic search for anatomic structures

2013 End of CO-ME, Transfer to SICAS
"SICAS Medical Image Repository" (SMIR)

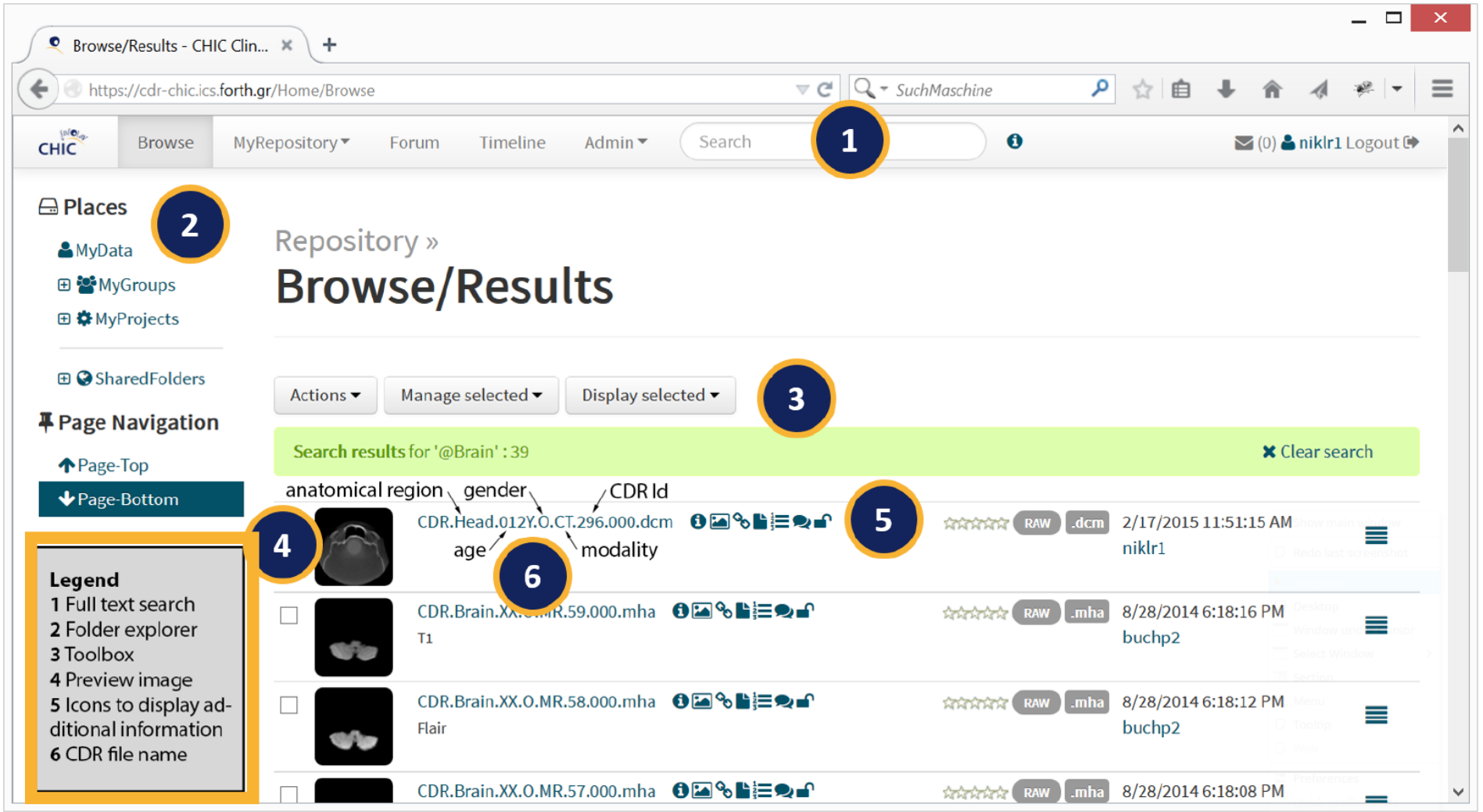
2013–now Co-development SICAS / CHIC-project
"Clinical Data Repository" (CDR)

- REST-API
- Support for (pre-)clinical data formats
- Enhanced semantic annotation & search

CHIC – Computation Horizons In Cancer

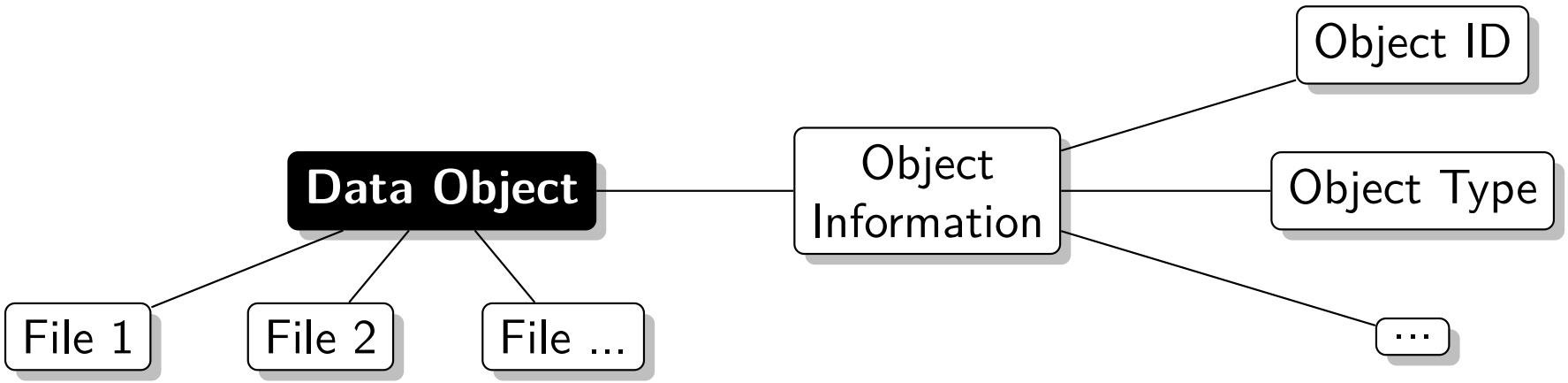
- EU FP7 "ICT Large-scale Integrating Project"
- Data repository for development of computational models in *in-silico* oncology

WEB Interface



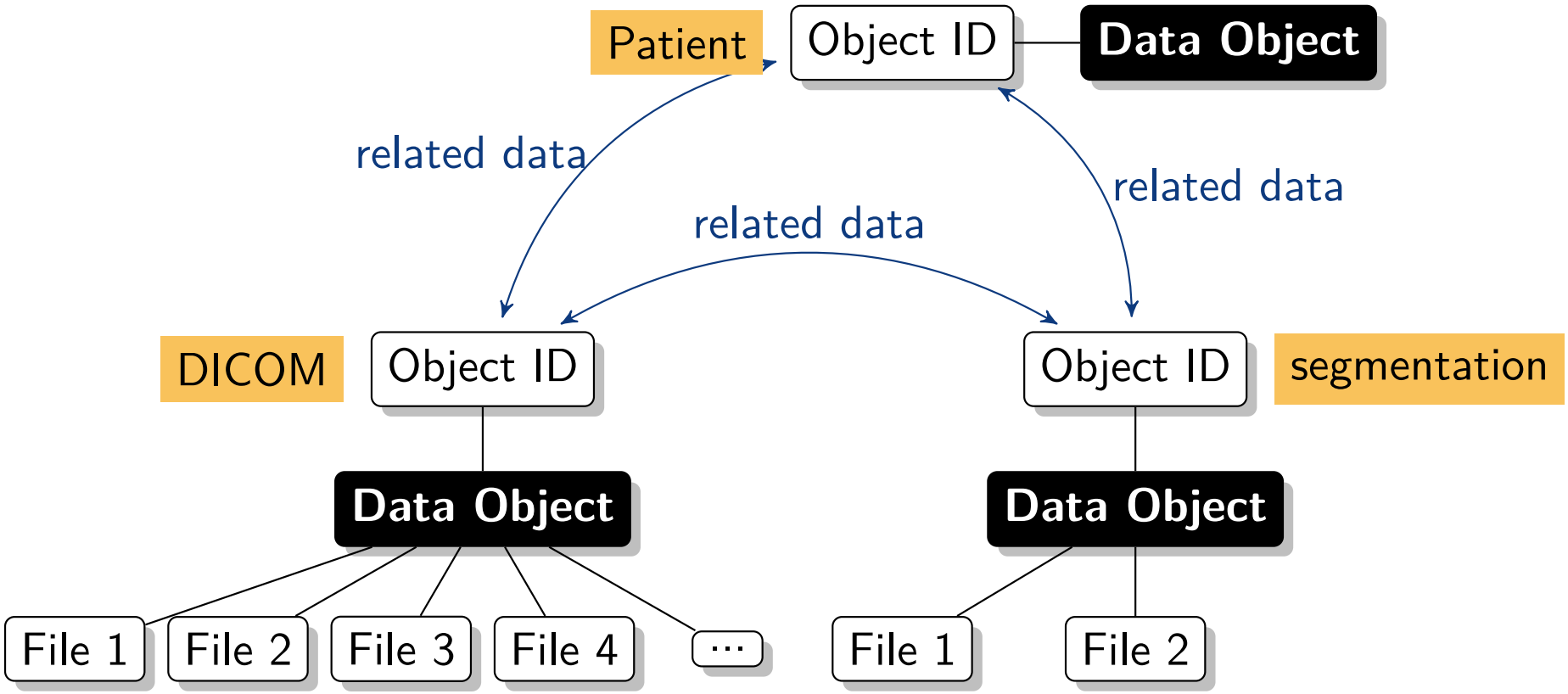
Object-based data model

- No fixed schema – everything is an "object"
- May contain multiple files (e.g. DICOM series, 2-file segmentation)
- Minimum set of additional object metadata



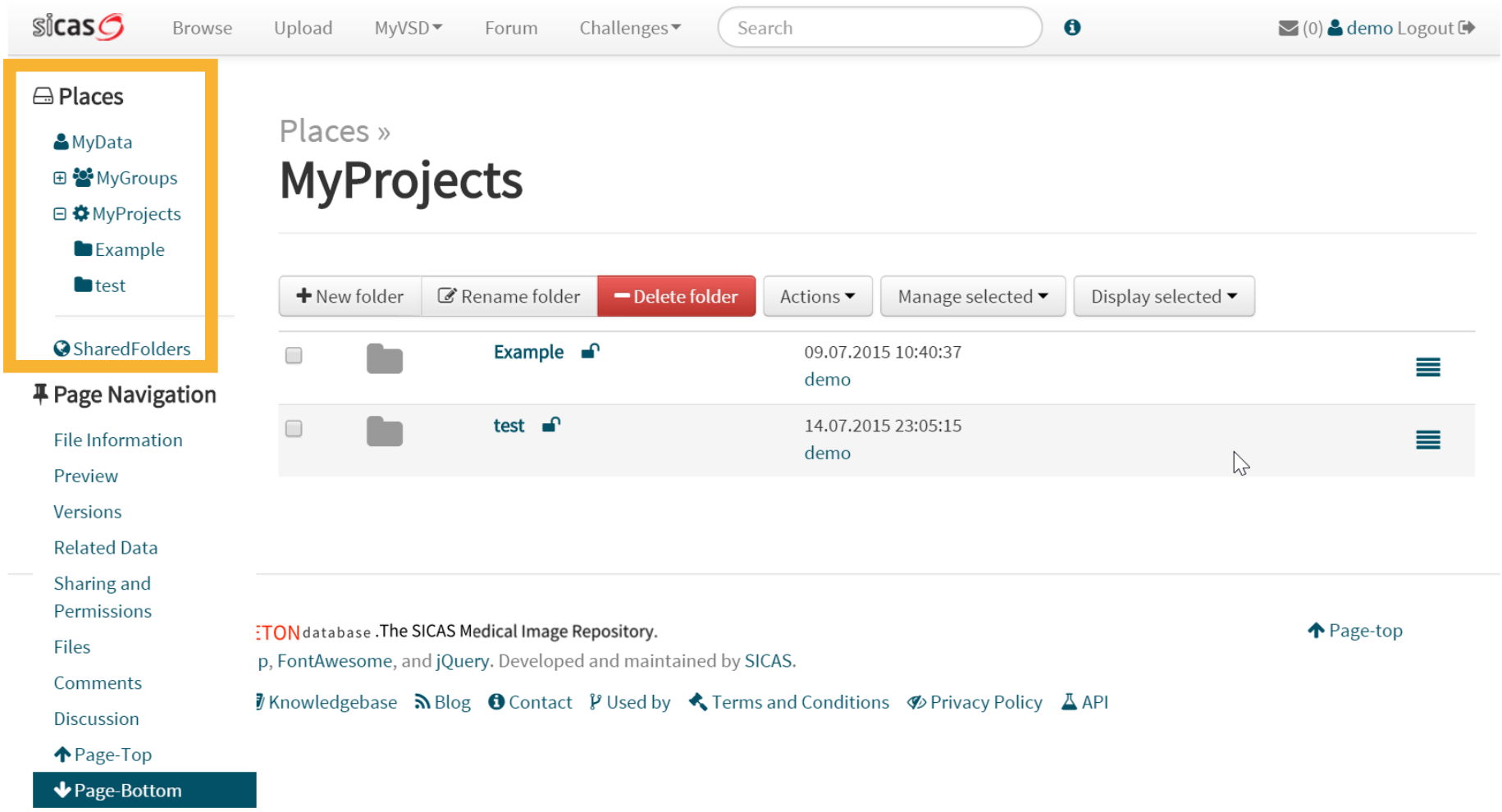
Related Objects

- Objects can be linked to other objects as “related data”
- Clusters intrinsically related information
- Same mechanism for grouping patient data



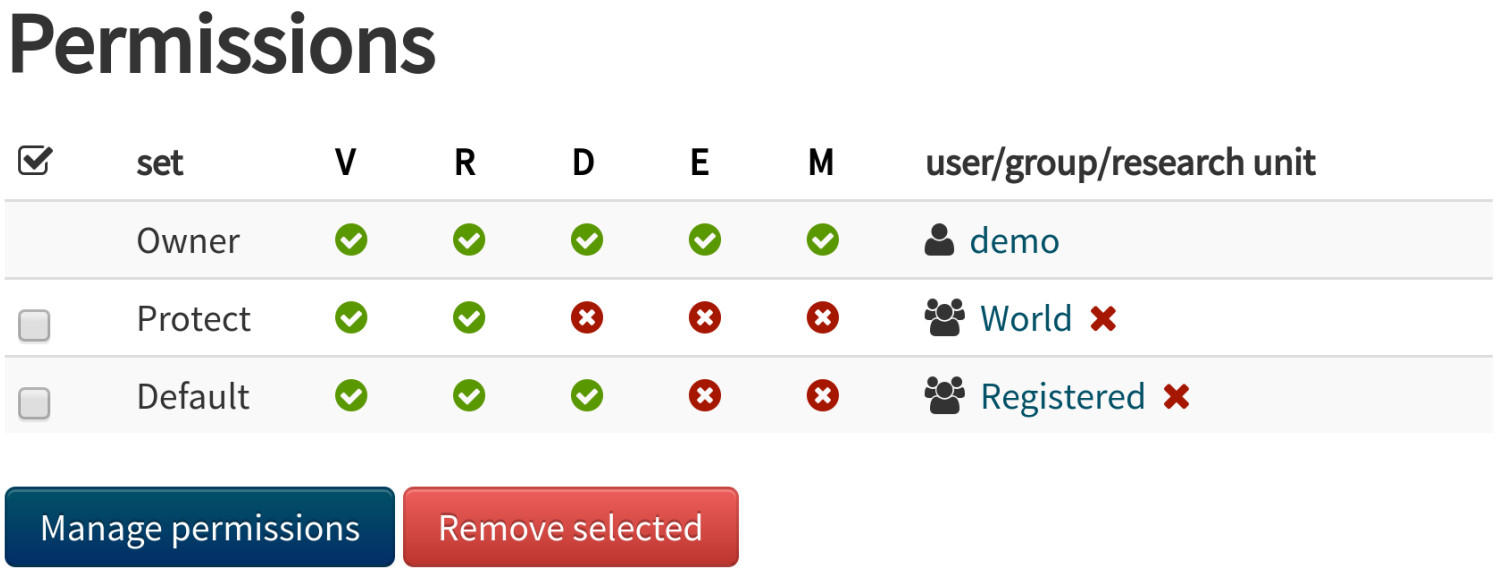
Custom Data Organisation

- User can group data in “folders”
- Custom organisation and sharing



Data Permissions

- Permissions can be set per object
- Fine-grained access attributes per user, group or research unit

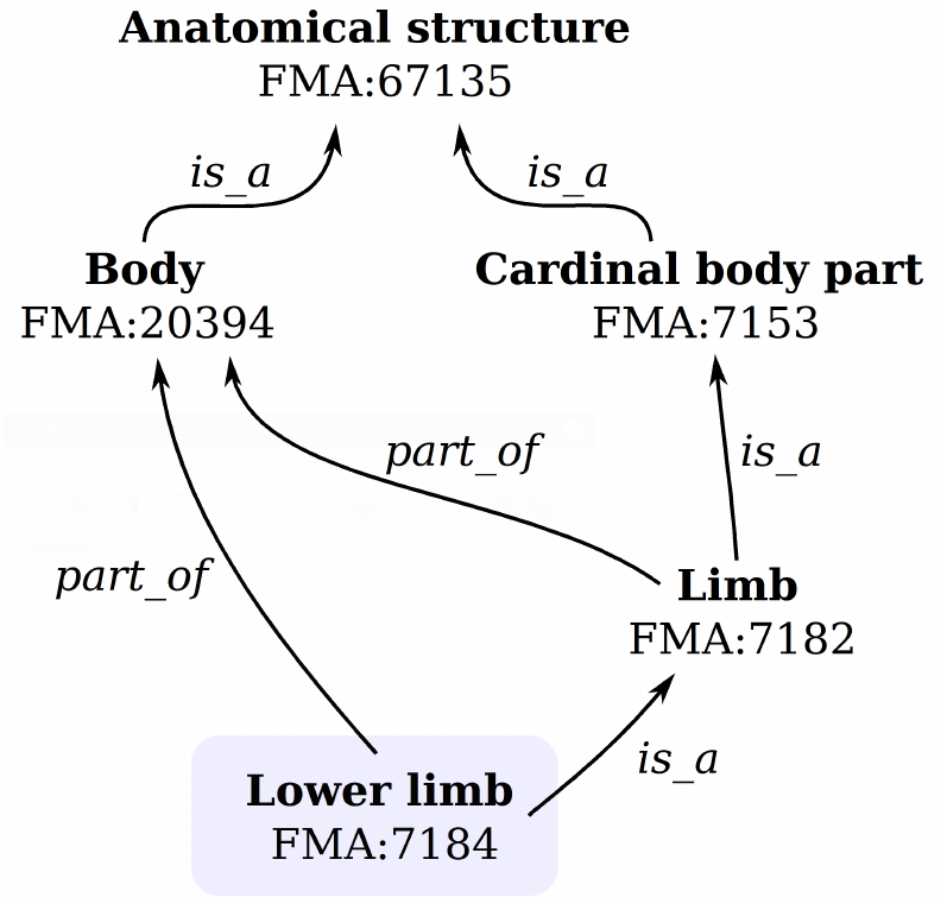


V – visit	See summary information
R – read	See all metadata information
D – download	Download data
E – edit	Edit data entry, i.e. metadata & files
M – manage	Manage data relations, access rights, etc.

Multiple Search Mechanisms

- Object ID
- Full text search in metadata

- Semantic search with ontology-based reasoning
  - Foundational Model of Anatmoy (FMA)
  - Other ontologies via RICORDO



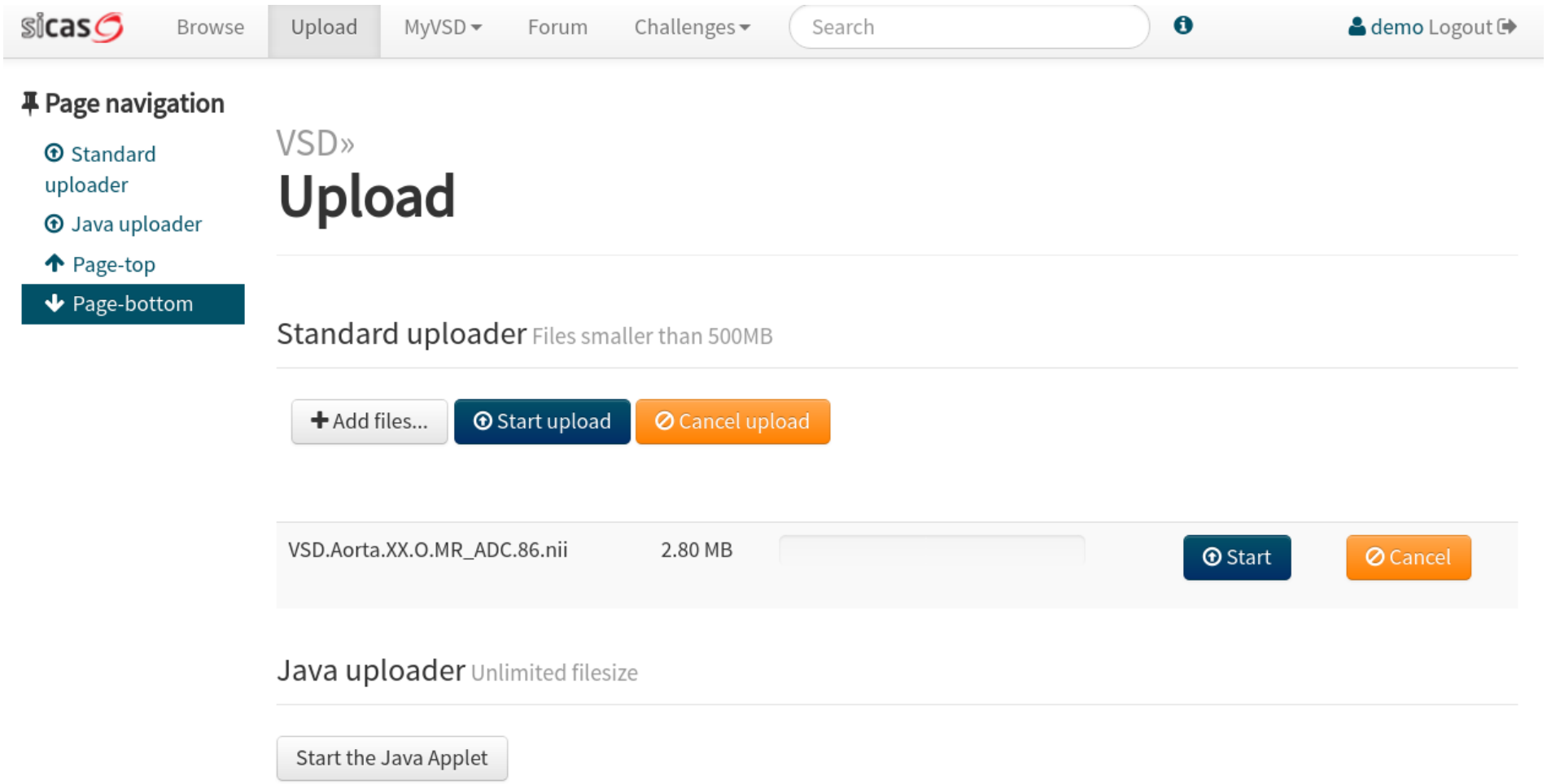
Browse/Results

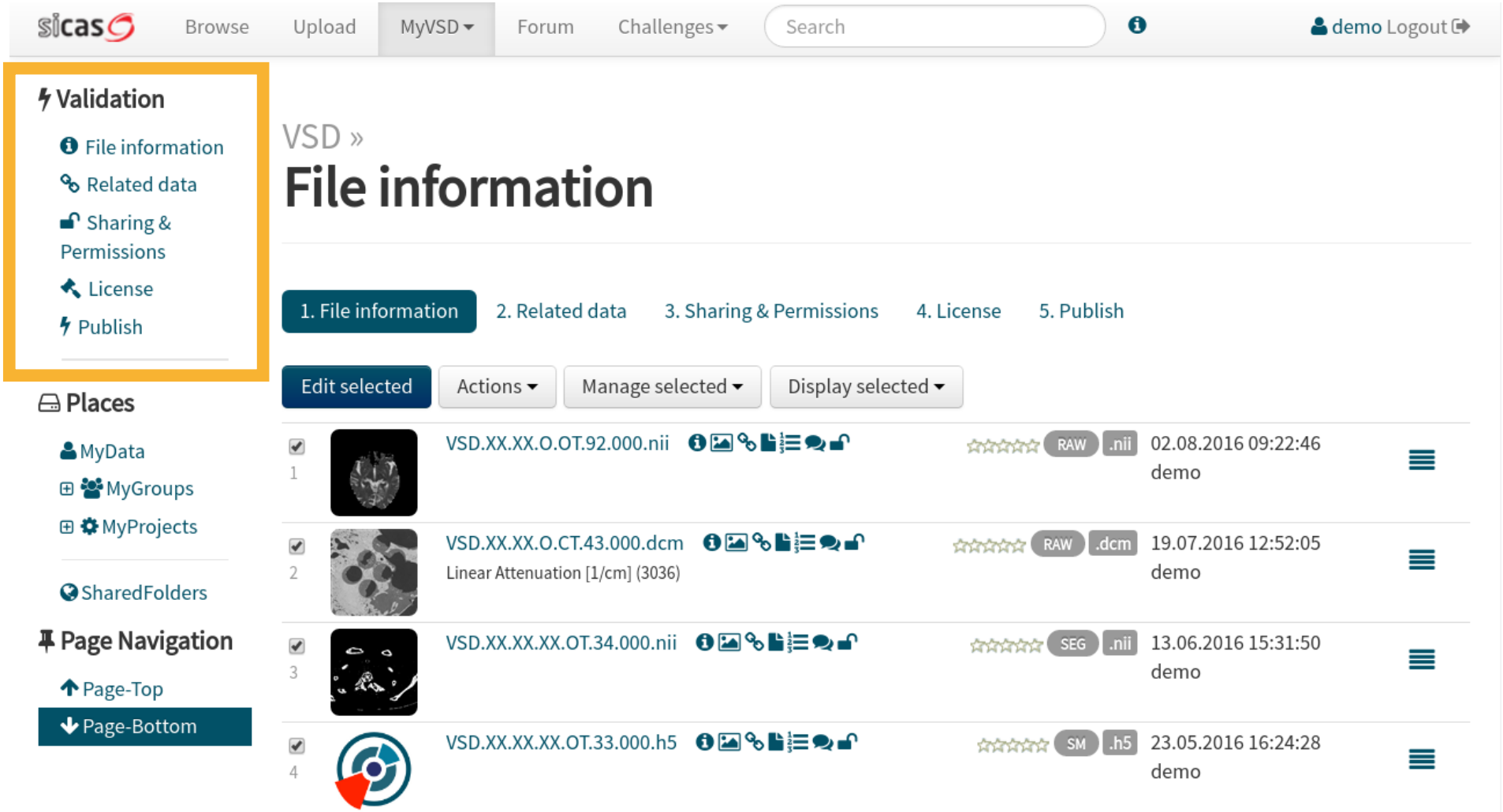
Actions ▾ | Manage selected ▾ | Display selected ▾ | 🔍

Search results for '@Femur' : 1893 ✖ Clear search

1		VSD.Femur.069Y.M.CT.2476.000.dcm INJECTION/Vol./CE/FC21/Q06/	RAW .dcm	13.07.2012 11:07:28 reyem1	☰
2		VSD.Femur.055Y.M.CT.2474.000.dcm INJECTION/Vol./CE/FC21/Q06/	RAW .dcm	13.07.2012 11:00:41 reyem1	☰
3		VSD.Femur.053Y.M.CT.2473.000.dcm MBINF	RAW .dcm	13.07.2012 10:54:24 reyem1	☰
4		VSD.Femur.XX.M.OT.1285.000.nii vsd 007	SEG .nii	11.06.2012 11:48:23 kistm1	☰
5		VSD.Femur.XX.M.OT.1283.000.nii VSD 005 Full body registered (Bern)	SEG .nii	11.06.2012 11:47:28 kistm1	☰
6		VSD.Femur.XX.M.OT.1281.000.nii VSD 004 Full body registered (Bern)	SEG .nii	11.06.2012 11:46:41 kistm1	☰
7		VSD.Femur.090Y.F.OT.1279.000.nii VSD 003 Full body registered (Bern)	SEG .nii	11.06.2012 11:45:49 kistm1	☰
8		VSD.Femur.078Y.F.OT.1277.000.nii from VSD 002	SEG .nii	11.06.2012 11:45:14 kistm1	☰
9		VSD.Femur.XX.XX.OT.1275.000.nii vsd 001	SEG .nii	11.06.2012 11:40:33 kistm1	☰
10		VSD.Lower_limb.064Y.F.CT.568.000.dcm lower leg, Unt.Extr.OS 1.25 B30s	RAW .dcm	28.02.2012 11:25:09 kistm1	☰

Data Upload


sicas
Browse
Upload
MyVSD
Forum
Challenges
Search
demo Logout
Page navigation
Standard uploader
Java uploader
Page-top
Page-bottom
VSD»
Upload
Standard uploader Files smaller than 500MB
Add files...
Start upload
Cancel upload
VSD.Aorta.XX.O.MR_ADC.86.nii
2.80 MB
Start
Cancel
Java uploader Unlimited filesize
Start the Java Applet

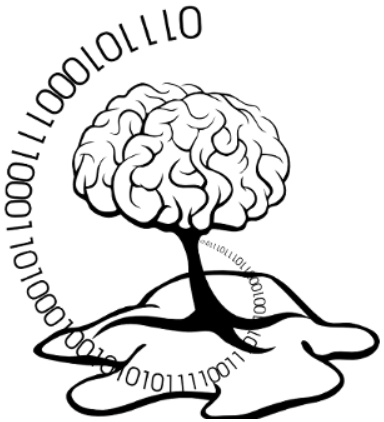

sicas
Browse
Upload
MyVSD
Forum
Challenges
Search
demo Logout
Validation
File information
Related data
Sharing & Permissions
License
Publish
Places
MyData
MyGroups
MyProjects
SharedFolders
Page Navigation
Page-Top
Page-Bottom
VSD »
File information
1. File information
2. Related data
3. Sharing & Permissions
4. License
5. Publish
Edit selected
Actions
Manage selected
Display selected
1
VSD.XX.XX.O.OT.92.000.nii
RAW
.nii
02.08.2016 09:22:46
demo
2
VSD.XX.XX.O.CT.43.000.dcm
Linear Attenuation [1/cm] (3036)
RAW
.dcm
19.07.2016 12:52:05
demo
3
VSD.XX.XX.XX.OT.34.000.nii
SEG
.nii
13.06.2016 15:31:50
demo
4
VSD.XX.XX.XX.OT.33.000.h5
SM
.h5
23.05.2016 16:24:28
demo

Benchmarking new algorithms

- Publish dataset & Task
- Evaluate submissions and rank participants
- Make raw data and results available to public

SHAPE

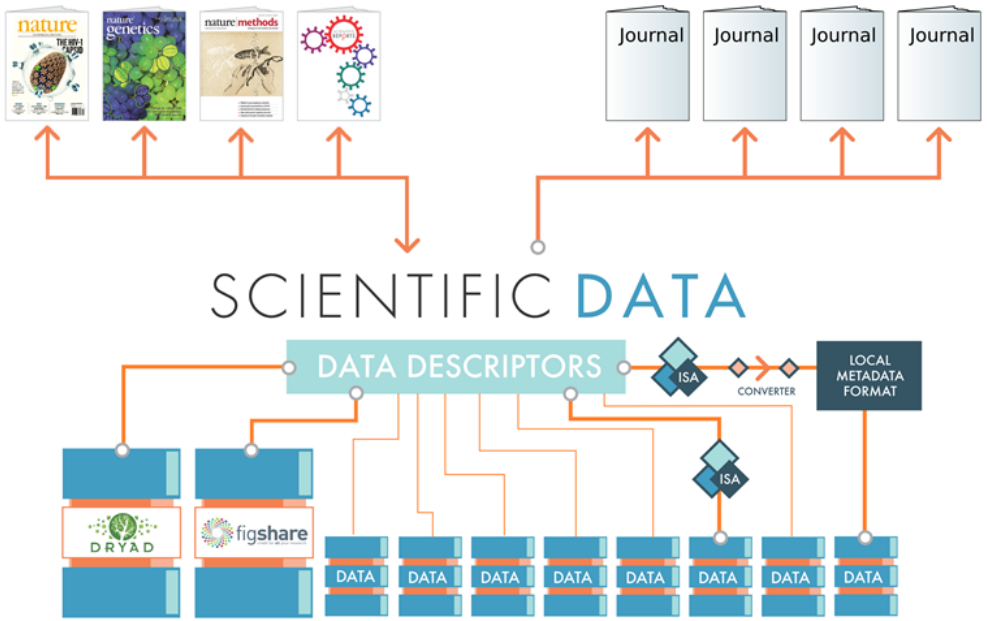
ISLES



BraTS

Open Access data for reproduceable research

- Storage for data supplementing published articles.
- Article describes dataset, data analysis and results.
- Data publicly available to enable reproduceability of research.



Source: Scientific Data Blog

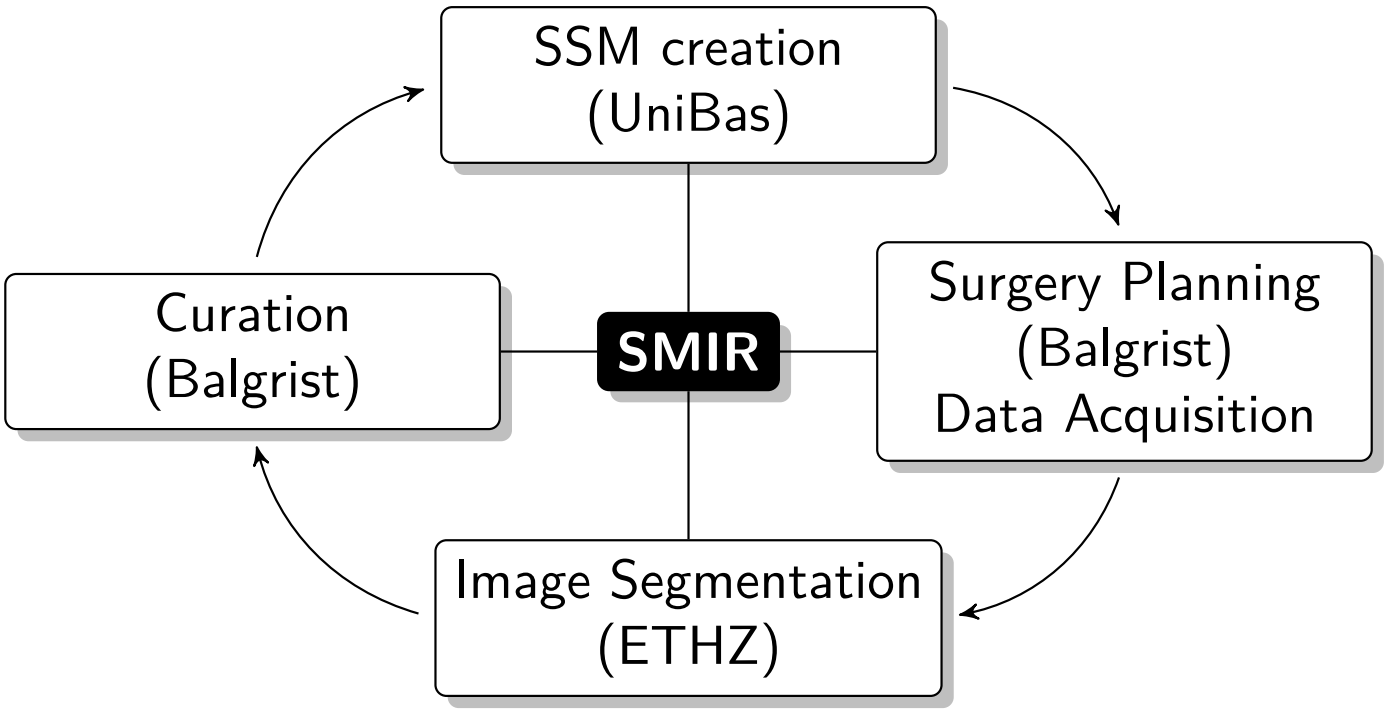
SICAS SMIR recommended by:

SCIENTIFIC DATA

PLOS

Statistical Shape Modeling for surgery planning

- Research data shared with project collaborators
- Data processing results added & linked to raw data
- Results used for planning medical intervention



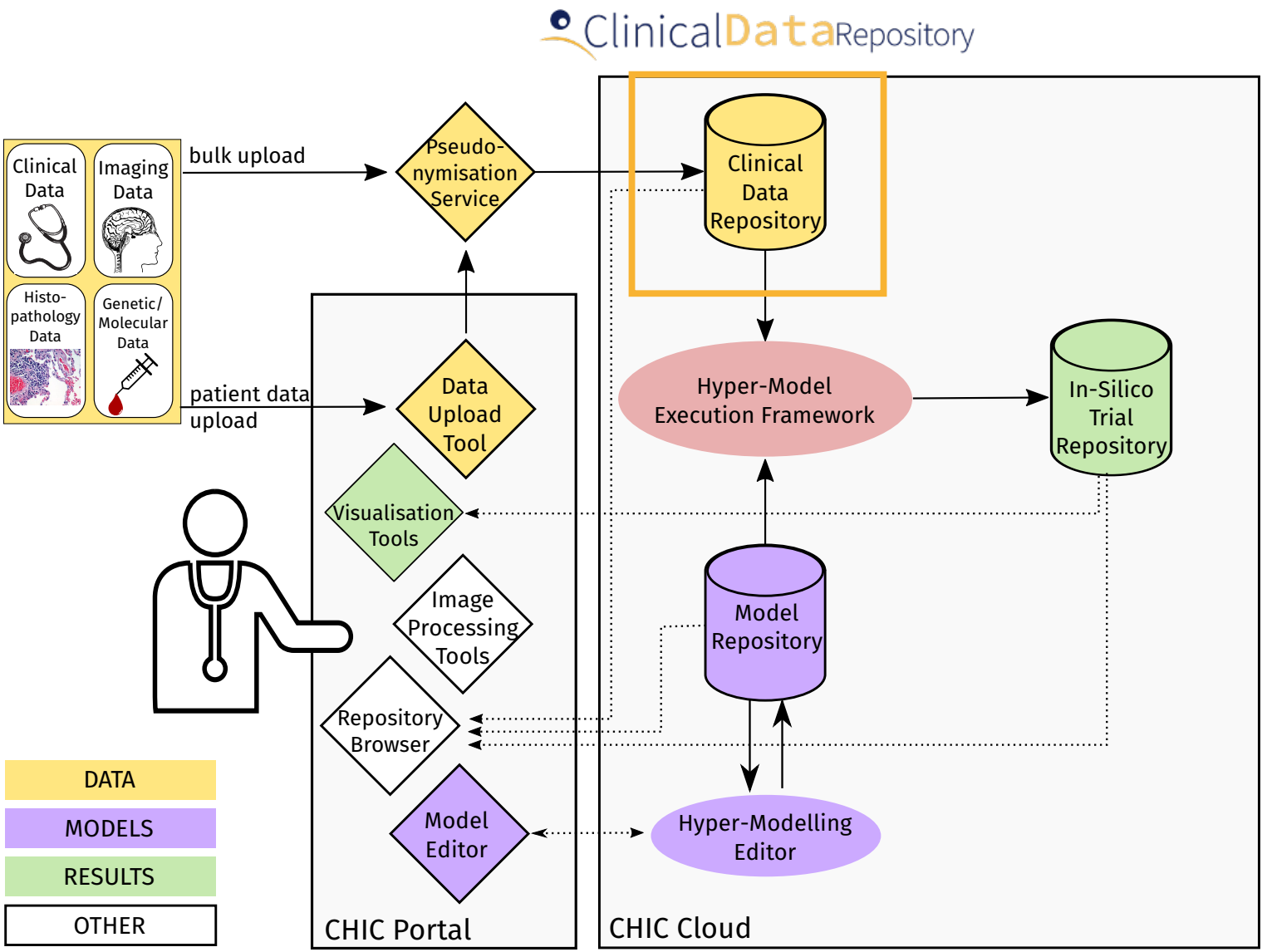
Data Repository for *In-Silico* Oncology

- *Clinical data* collected from multiple clinical partners.
- Shared for *research* with collaborators.
- Repository integrated into larger system
  - CHIC Data Protection Framework
  - CHIC Security Framework

ClinicalDataRepository

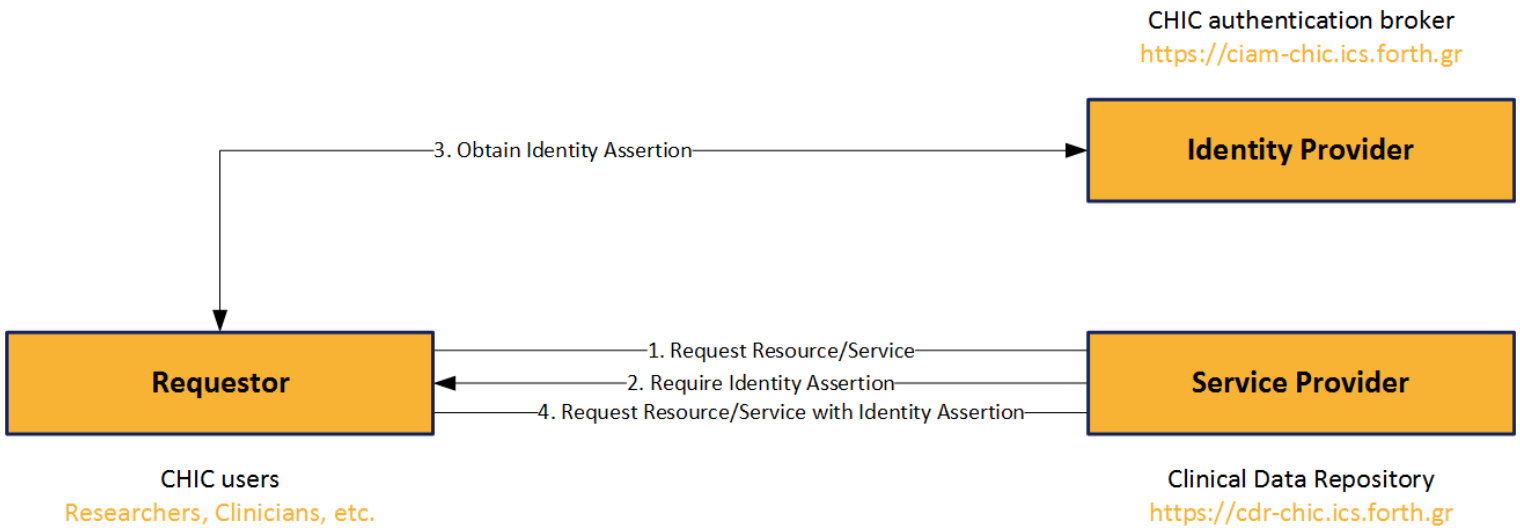
Contents

CDR in CHIC


ClinicalDataRepository
Clinical Data
Imaging Data
Histo-pathology Data
Genetic/ Molecular Data
bulk upload
patient data upload
Pseudo-nymisation Service
Clinical Data Repository
Data Upload Tool
Hyper-Model Execution Framework
In-Silico Trial Repository
Visualisation Tools
Image Processing Tools
Model Repository
Repository Browser
Model Editor
Hyper-Modelling Editor
DATA
MODELS
RESULTS
OTHER
CHIC Portal
CHIC Cloud

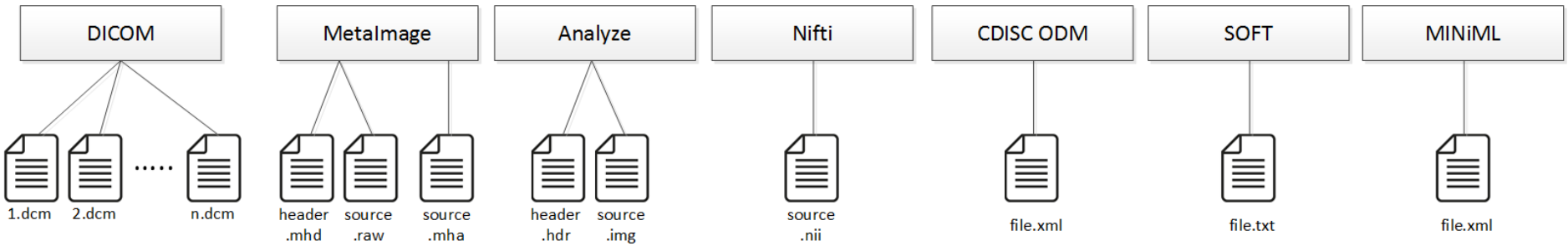
Deployment on CHIC Infrastructure

- Clinical Data Repository (CDR) running on CHIC cloud (FORTH)
- CDR integrated into CHIC security framework
  - Brokered authentication
  - Single Sign-On (SSO) through website (Shibboleth) and REST API (SAML)

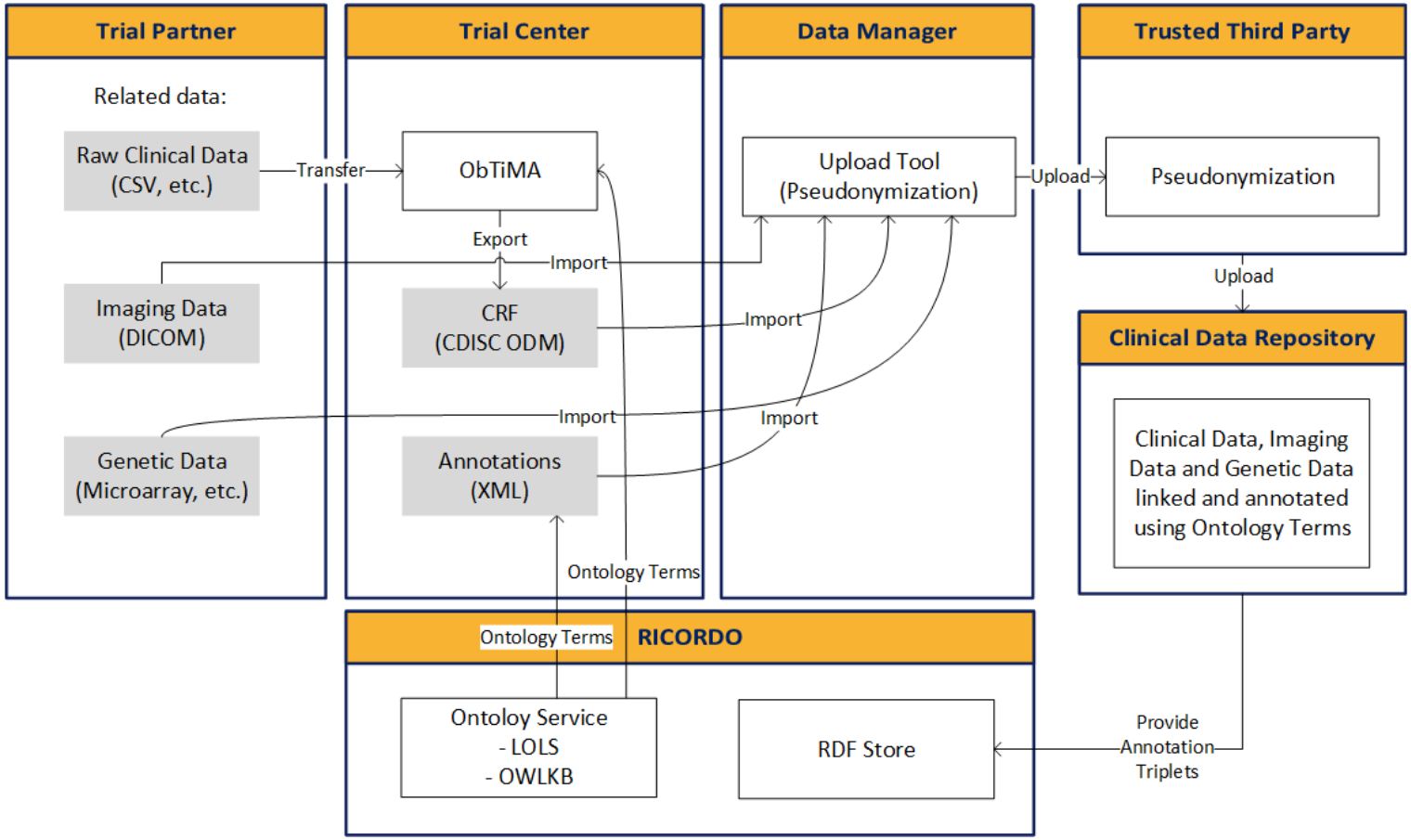


Data in CHIC-DCR

Data Type	Standard Format	Source
Clinical Data	CDISC ODM	ObTiMA
Imaging Data	DICOM, MetaImage, Analyze,NIfTI	local PACS
Genetic / Molecular Data	MINiML	platform specific
Histopathology	CSV, JPEG	platform specific



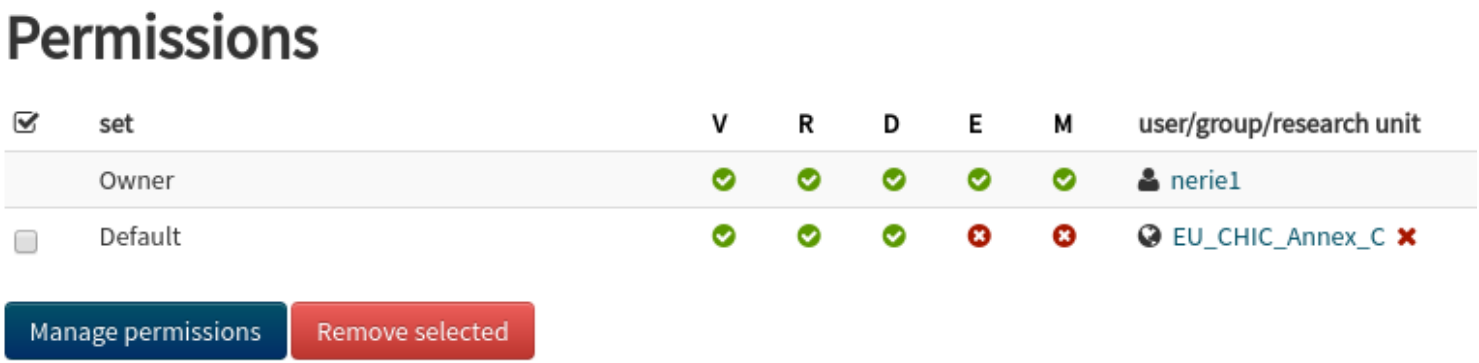
Data Upload Workflow



- Trial Center collects patient data from trial partners
- Trial Center uploads data to "Trusted Third Party" (TTP) (1st pseudonymization)
- TTP uploads data to CDR (2nd pseudonymization)

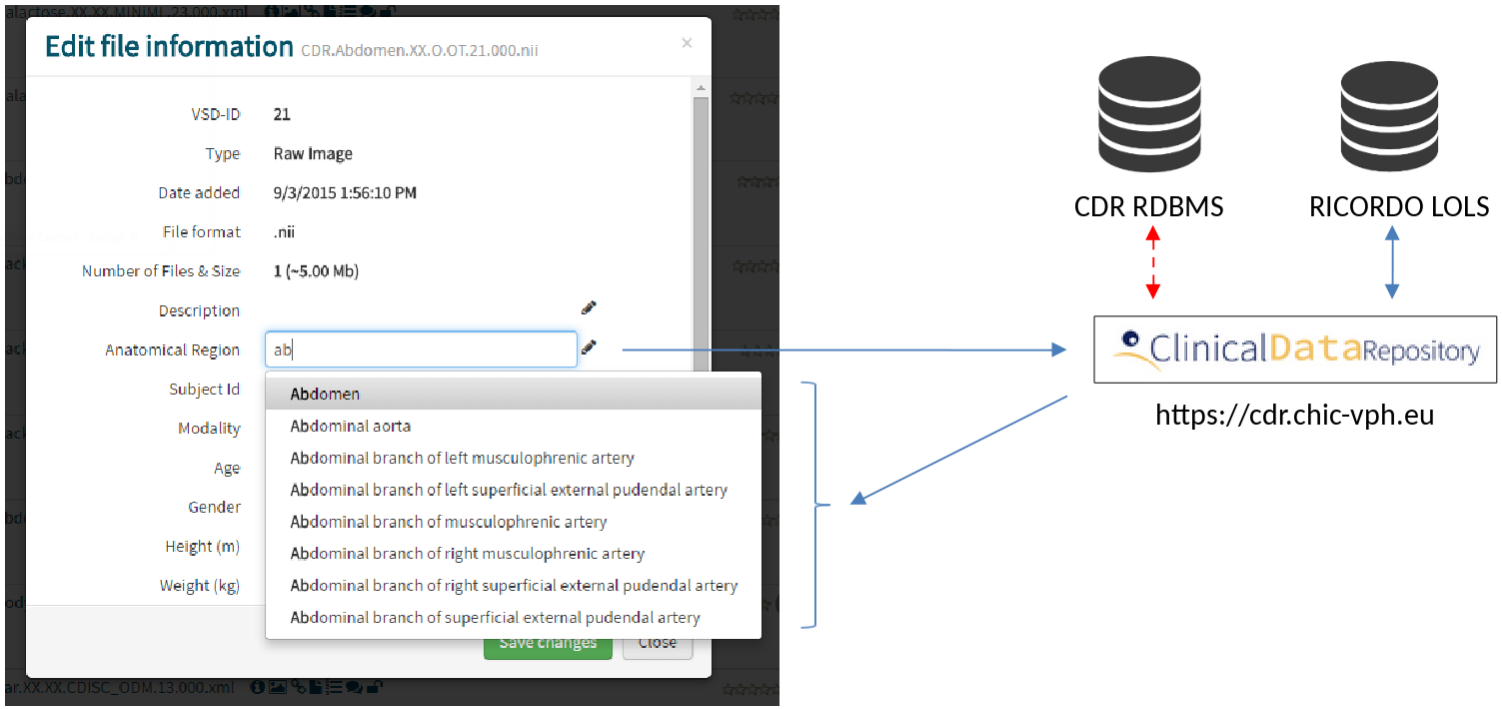
Data Protection & Access

- CHIC infrastructure embedded in CHIC Data Protection Framework
  - Patients' informed consent and/or ethics board approval
  - Data de-facto anonymised (2 pseudonymisation steps)
  - Trusted Third Party (TTP) holds key for 2nd pseudonymisation step
- Data access and exchange by users subject to Annex C: "Access authentication form"
  Contract between CHIC End User and Center for Data Protection (CDP)



Semantic Annotation

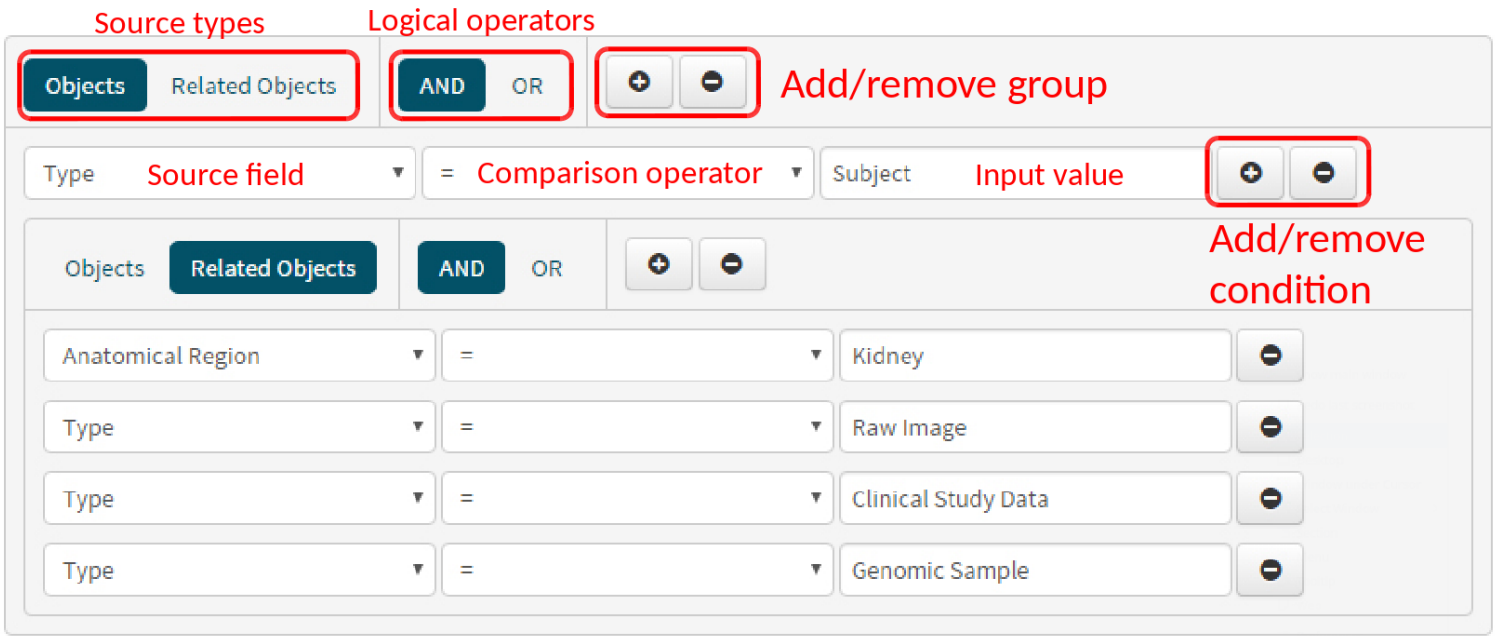
- Improved semantic annotation, auto-complete for multiple ontologies
- Using RICORDO `http://www.ricordo.eu/` service



Advanced Meta-Data Search

- Identification of "complete" data sets for research

  Find patients for whom we have imaging, clinical and miRNA nephroblastoma data



Data Repositories for *In-Silico* Oncology

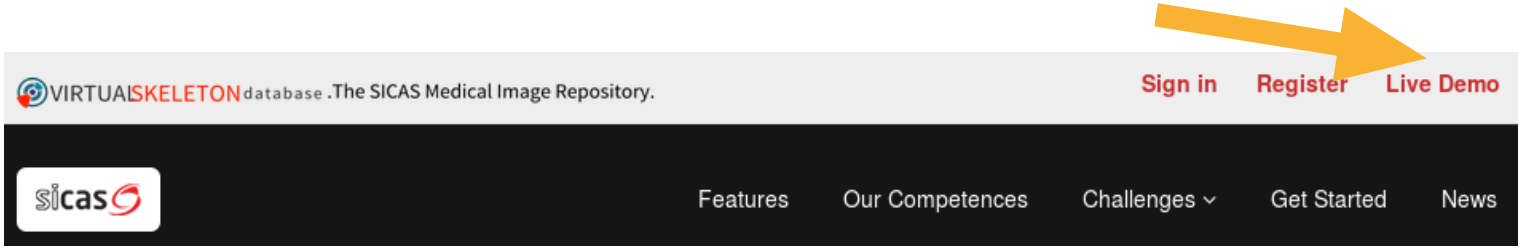
- Must-Haves
  - Legal framework for data provision, use and sharing
  - Programmatic access (including authentication)
  - Auditing
  - Strong search capabilities
- Challenges
  - Data curation
  - Data granularity
  - Bridging research and treatment scenarios

Contents

Summary

- CHIC Clinical Data Repository is branch of SMIR: `https://www.smir.ch/`
  - Fine-grained access control
  - Flexible data structure for customisable organisation
  - Standardized formats & semantic annotation
- Adapted for in-silico trials:
  - Semantic search enables data discovery
  - REST-based API allows integration in web-service architecture
  - Auditing

- SMIR for your research?
  - Project-based collaboration
  - Demo at `https://www.smir.ch/`



Funding

This project has received funding from the European Union Seventh Framework Programme for research, technological development and demonstration under grant agreement n° 600841.

THANK YOU!

SMIR for your research?

Contact us at:

- https://www.smir.ch/
- philippe.buechler@istb.unibe.ch