



BioMedBridges



Common Challenges Across Scientific Disciplines

Laurence Field

CERN

18th November 2013

- Facilities, resources and related services
 - Used by the scientific community
 - For research in their respective fields
- Examples
 - large-scale research installations
 - Libraries, databases and archives
 - Integrated research installations
 - Communication networks
 - Distributed computing facilities
 - Data infrastructure
 - Observatories
 - Centres of competence
- May be single-sited, distributed or virtual



ESFRI



- Develop the integration of RIs in Europe
 - Addresses the fragmentation of policies
 - A coherent and strategy-led approach
 - Facilitates multilateral initiatives
 - Strengthen their international outreach
 - Leading to better use and development
- Nominated by the Research Ministers
 - Of the Member and Associate Countries
 - A representative of the Commission
 - A joint vision and common strategy
- The Roadmap
 - First published in 2006
 - Updated in 2008 and 2010

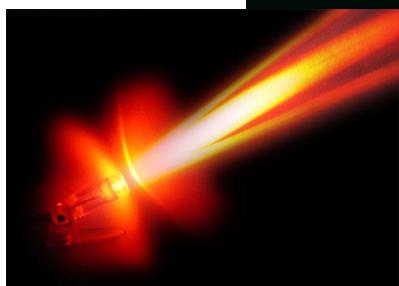
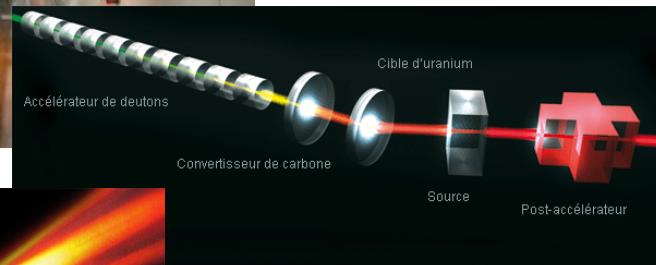


ESFRI

Social Sc. & Hum. (5)	Life Sciences (10)		Environmental Sciences (10)		Material and Analytical Facilities (6)	Physics and Astronomy (11)		Energy (4)	e-Infra-structures (1)
SHARE	BBMRI	ELIXIR	ICOS	EURO-ARGO	EUROFEL	ELI	TIARA	ECCSEL	PRACE
European Social Survey	ECRIN	INFRA FRONTIE	LIFEWATCH	IAGOC	EMFL	PRINS	CTA	JHP	
CESSD	INSTRUCT	EATRIS	EMSO	EPOS	European XFEL	SPIRE	SKA	IFMIF	
CLARIN	EU-OPEN SCREEN	EMBRC	SIAEOS	EISCAT_3D	ESRF Upgr	E-ELT	FAIR	HIPER	
DARIAH	Euro BioImaging	ERINHA BSL4 Lab	COPAL	AURORA BOREAL	NEUTRON ESS	KM3NeT	ILC-HIGRADE		
					ILL20/20 Upgr	SLHC			



Physics



Astronomy



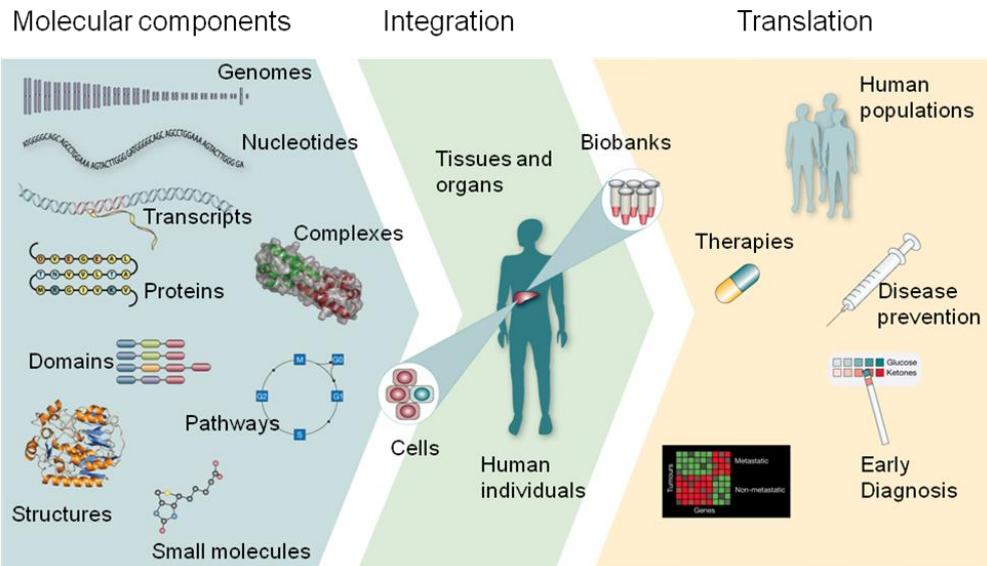
Analytical facilities





Objectives:

- Computational ‘data and service’ bridges between the BMS RIIs
- Interoperability between data and services in the biological, medical, translational and clinical domains (semantic integration)
- Link basic biological research and data to clinical research and associated data
- Connect existing data (no new data generated within the project)
- By connecting it “at a click”, make data easily discoverable for researchers



Data standards

aim amplitudes biomart bridg cdash
dicom dictionaries clinical commercial
data emdb envo file
finnish genomic standards consortium gff gic h7 lcd10 kegg kogg meddra miami
mixs mmcif mp mp ontology national nmr-star obsrv-om odm omim
ontology open geospatial consortium porom protein 3d coordinates
protocol sdtrm send services snomed ssh structural structural structural amplitudes transfer web xgap
xml

18s 23s 28s abundance amplitudes annotations biological
biosamples chemical cryo-preserved ct
describing em est etc expression gene genome
genotypes geo-referenced imaging medical mets metadata
metagenome microscope molecular mouse mri omics pathology
phenotypes preparation products protein purification qtlis
records registries related ribosomal samples sector sequence spec
structures study-level tomograms transcriptome

Types of data

aim arb bam bedgraph biomart cdisc cif CSV dcm dicom
fasta fastq files formatvcf gcdml genbank hkl jpeg
embl emdb images ins json lab lincs miami mysql mzdata mzml mzxml odm openclinica templates
pathology pdb pdf proprietary format sca scanner-specific sdtrm
tiff jpg tsv txt vcf xml

Data formats

The ten ESFRI BMS research infrastructures in BioMed Bridges cover:

- biobanks
- bioinformatics
- translational research
- clinical trials
- structural biology
- mouse biology
- imaging
- marine resources
- highly contagious agents
- chemical biology

- Enable multidisciplinary Environmental Scientists to access, study and correlate data from multiple domains for *system-level* research by providing solutions and guidelines for the RIs common needs.
- Main products:
 - Reference Model
 - Data Discovery tool
 - Data Processing environment





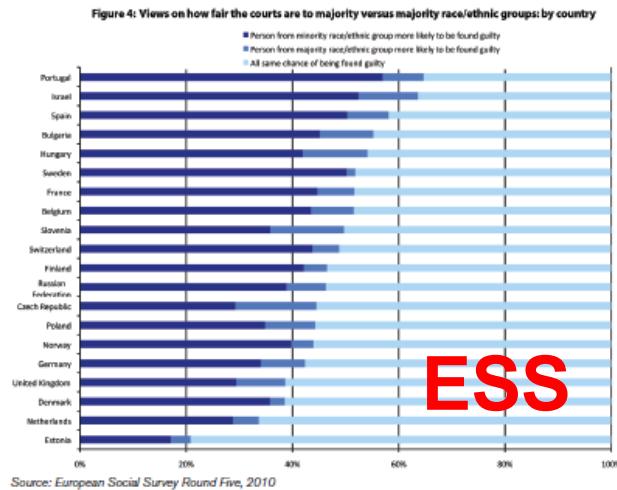
DARIAH

The Humanities Project

maintaining languages, minds and societies



sensing state of minds,
cultures and societies



ESS





Common Challenges

	CRISP	ENVRI	DASISH	BioMedBridges
Data identity				
Data continuum				
Software identity				
Concept identity				
User identity management				
Common attribute scheme				
Common data standards and formats				
Service discovery				
Service market places				
Integrated data access and discovery				
Data storage facilities				
Data curation				
Privacy and security				
Dynamic data management				
User community body				
Semantic annotations and bridging				
Reference models				
Education & training				

User Community Body

- Relevance of an e-infrastructure
 - Defined by its users and the scientific research that it supports
- Requirements of the researcher must drive the development
 - European e-infrastructure is growing
 - Important that user communities are able to voice requirements
 - Help drive the direction of their evolution.
- European e-infrastructure providers need requirements
 - From a wide variety of user communities
 - Not necessarily the end users themselves
 - Institute or project that supports
- For individual research infrastructures
 - Cooperating infrastructures in a scientific domain
 - Beneficial to exchange experiences with the user communities
- Many existing initiatives for user engagement
 - Applying many different strategies
 - Remain highly fragmented
 - The user forum should be seen as a complementary