ESFRI and e-Infrastructures

eIRG Workshop Uppsala 14-15 October 2009

Juni Palmgren, Stockholm

Outline:

- ESFRI eIWG 2008; TWG 2009
- ESFRI Roadmap 2008: Challenges for e-Infra
  - Examples
- A structured process 2009 and beyond?

ESFRI European Strategy Forum on Research Infrastructures

Terms of Reference for a Transverse Working Group Dealing with e-infrastructures in the context of the update of the ESFRI Roadmap in 2008

- Critical to the success of every RI is the supporting electronic infrastructure (e-Infrastructure). Indeed, this will be a crucial factor for the majority of projects. As such, ESFRI agreed the setting up of a "Transverse Working Group" dealing with the e-Infrastructure aspect of proposals put forward for the update of the ESFRI Roadmap.
- The **"e-Infrastructure Transverse Working Group" (e-IWG)** shall work in parallel and in cooperation with the 4 existing Roadmap Working Groups (RWGs) (BMS, PSE, ENV and SSH).
- It is understood that the e-IWG should not lead the other RWGs but rather provide a common perspective on e-Infrastructure needs.

eIWG 2007-08

- 26 proposals reviewed
- 6 of these on the updated roadmap (2/6 included scored mature for e-Infra; 2/18 not included scored mature for e-Infra; 2 missing)

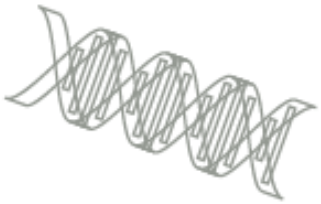
- General conclusion:

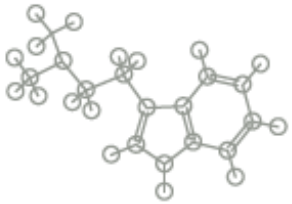
  The overall feeling of the meeting is that the level of proposals is not very high, as far as e-infrastructures are concerned. *Only a few of the proposal are considered as mature in this field. Most of them are considered at best as emerging.* Many of the proposals do not really identify the content of e-infrastructures and the potential benefit for them. Many think that access to Internet is sufficient to fulfill their needs.

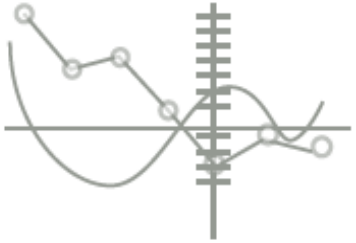
>The Roadmap: the landscape and its projects

>Social sciences and Humanities 17

>Environmental Sciences 25







ESFRI European Strategy Forum on Research Infrastructures

Terms of Reference
Thematic Working Groups

The ESFRI Roadmap for Research Infrastructures, published in 2006 and updated in 2008, is a vital policy document and paves the way for the planning, implementation and upgrading of RIs for the coming decades. ESFRI is committed to stimulate the implementation of these facilities and update this document as the need arises.

The work of ESFRI **Thematic Working Groups** (TWGs) underpins this strategic process.

Thematic Domains

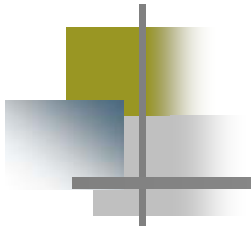
The following scientific domains have so far* been identified and shall have their own TWG:

- Biological and Medical Sciences (BMS)
- Energy (ENE)
- Environmental Sciences (ENV)
- Engineering Sciences and Physics (ESP)
- Social Sciences and Humanities (SSH)
- E-Infrastructures (e-IWG)

Frozen and task taken up by eIRG (as of Oct 9, 2009)

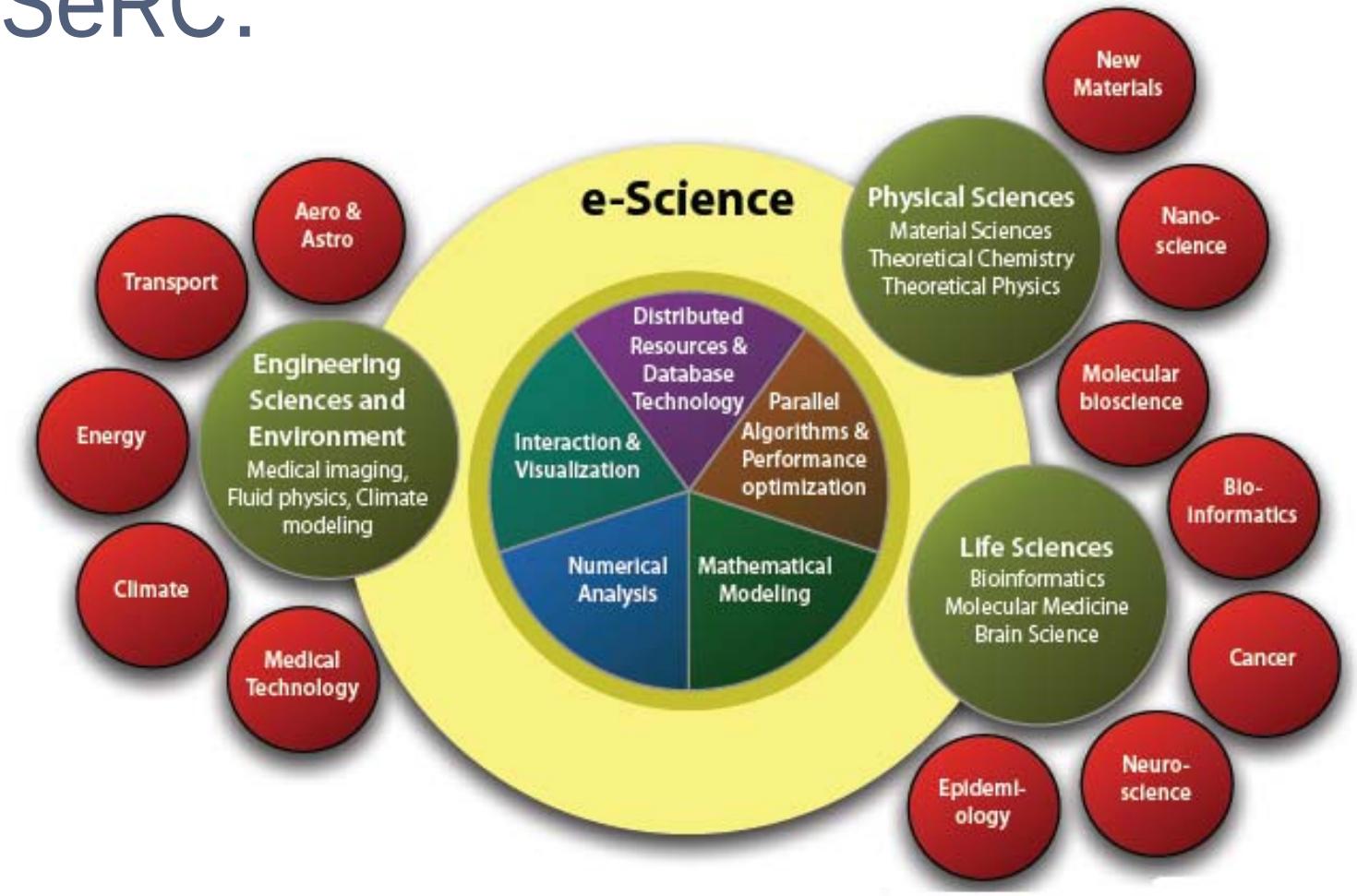
From HPC centers to e-Science environments/communities

- e-Science environments
  - Service orientation
  - Integration of e-Infra
  - Interfaces
  - Virtualisation
  - etc.
- Data infrastructures
- High Performance Computing
- Visualisation tools
- etc.



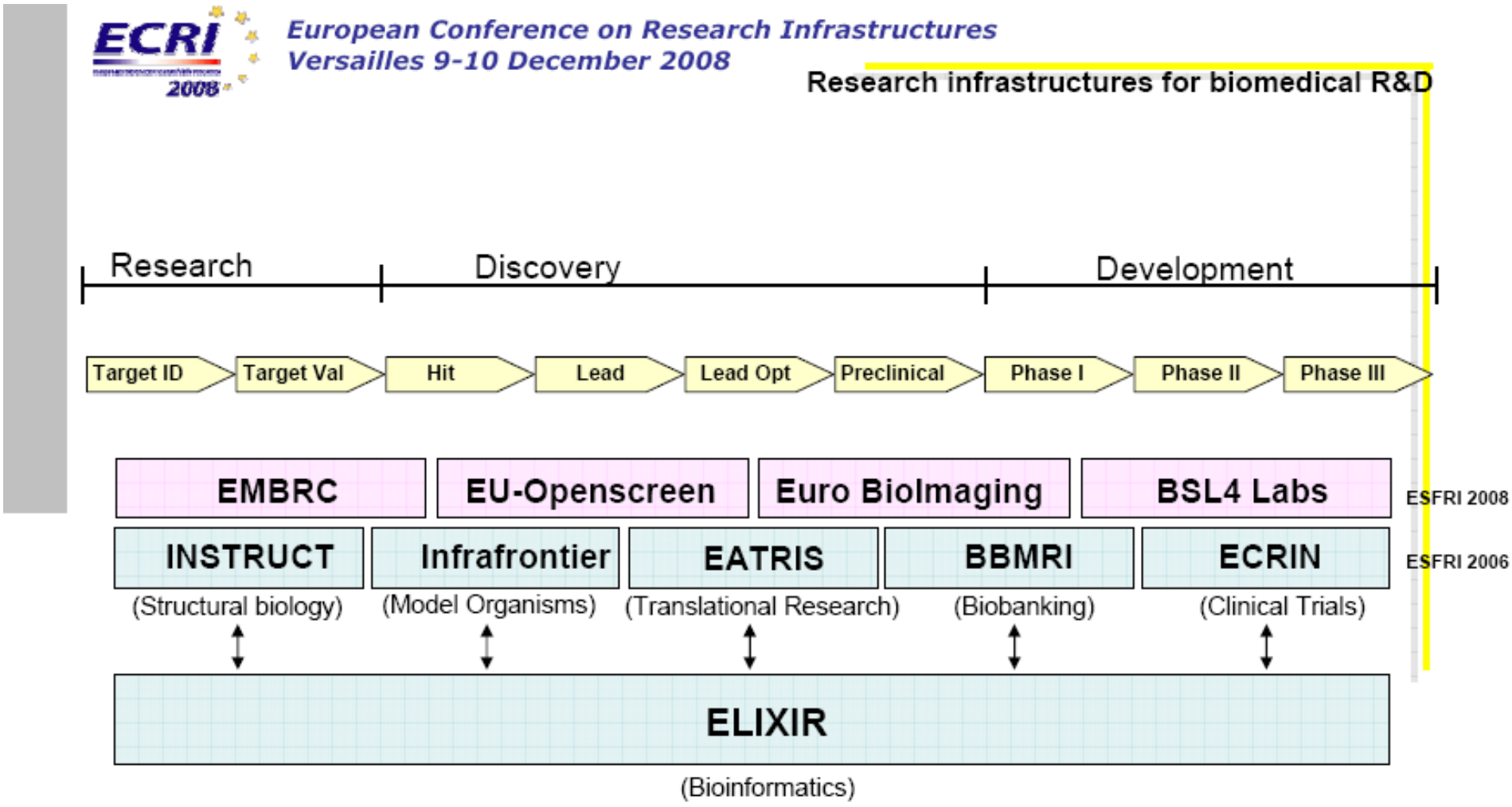
HOW?

Swedish e-Science Research Centre SeRC:



Swedish government strategic research areas 2009:
e-Science: SeRC is joint for KTH, SU, LiU, KI, with start 2010

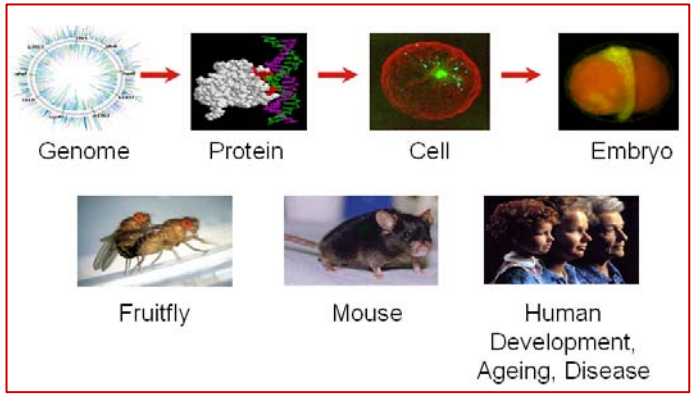
Iain Mattai, EMBL:



ESFRI Roadmap Update 2008: European Marine Biology Resource Centre (EMBRC)
EU-Openscreen (screening platform for chemical biology, Euro BioImaging, High Security Laboratories (BSL4 labs)

Examples of data from BMS RI

- Genes, genomes, proteomes, gene expression data
- Molecular structures
- Human tissue samples
- Small chemical molecules and their biological activity
- Imaging: confocal microscopy and medical imaging data
- Clinical trial data
- Model animal mutants and their phenotypes
- Translational research data
- Medical (patient) data



New types of data are continuously emerging due to technology development!

Torrents of new biological data!

Annotation of sequence variation catalogue
is of crucial importance

- 2008 genome sequencing: 1.5 petabytes data in hundreds of centers in dozens of countries with two Tier 0 sites (EMBL-EBI and NCBI)
- New sequencing technology 5-10 fold capacity
- Bioinformatics requirements for large scale biology are growing faster than computational capacity and transfer methods – new and smarter solutions needed!

Note:

- LHC produces 15 petabytes per year
- The LHC grid is 140 computer centres in 33 countries centered on CERN as Tier 0.

Process for 2009 and beyond?

Possible dialogue between RI & eIRG:

- Invitation from eIRG
- RI (RI domain) self assessment of e-Infra needs
  - Organisation
  - Access
  - Service to scientists
  - Other..
- Hearing/workshop
- Plan for e-Infra needs and structure
- Plan for follow-up