The e-Infrastructure needs of ESFRI projects

Yannis Ioannidis

MaDgIK Lab

University of Athens   &   ATHENA Research Center

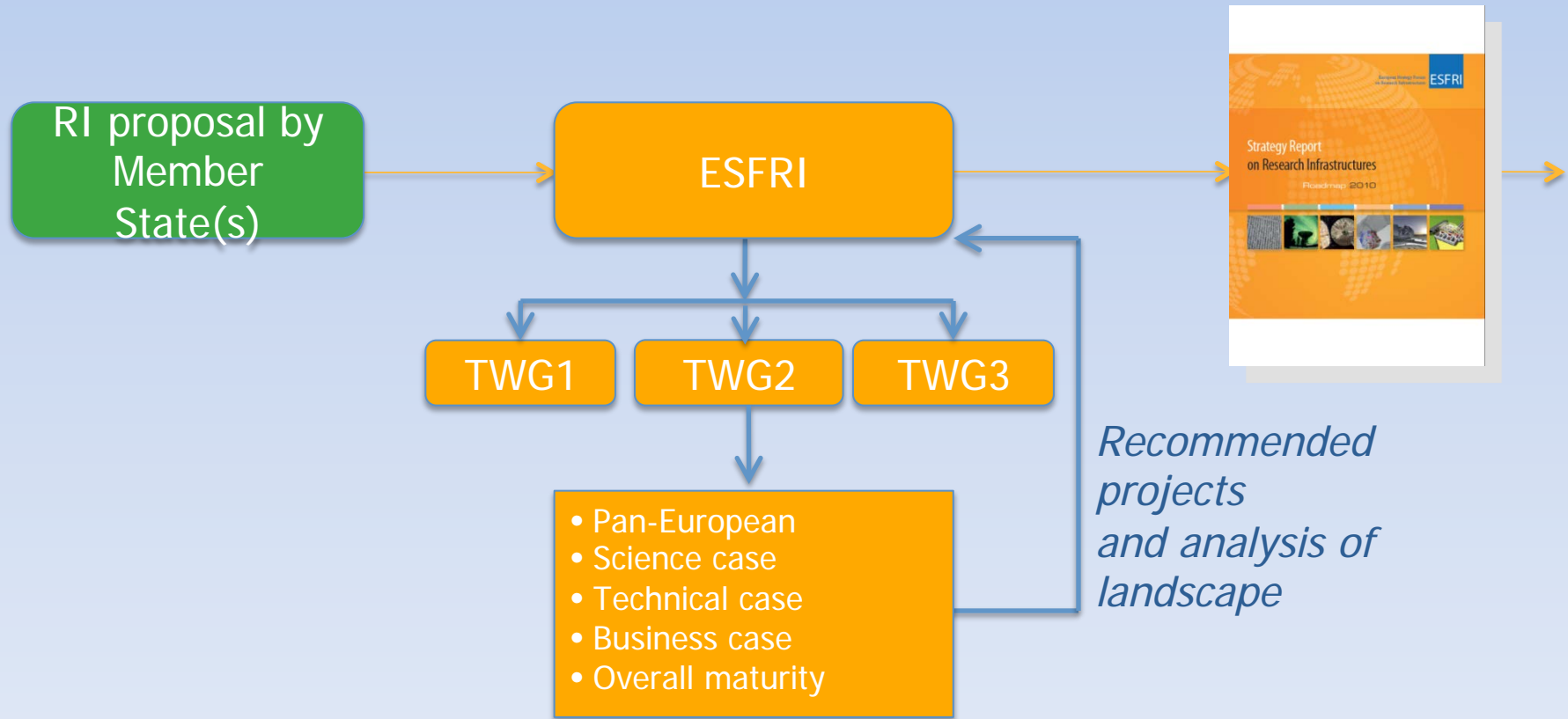# ESFRI: European Strategy Forum on Research Infrastructures

- ◦ Set up by the EU Council of Research Ministers in 2002

- ◦ Brings together representatives of Ministers of the 27 Member States, 10 Associated States, and of the EC

- ◦ Supports a coherent and strategy-led approach to policy making on Research Infrastructures

- ◦ Mandate to develop a Roadmap (2006) and its updates (2008, 2010)

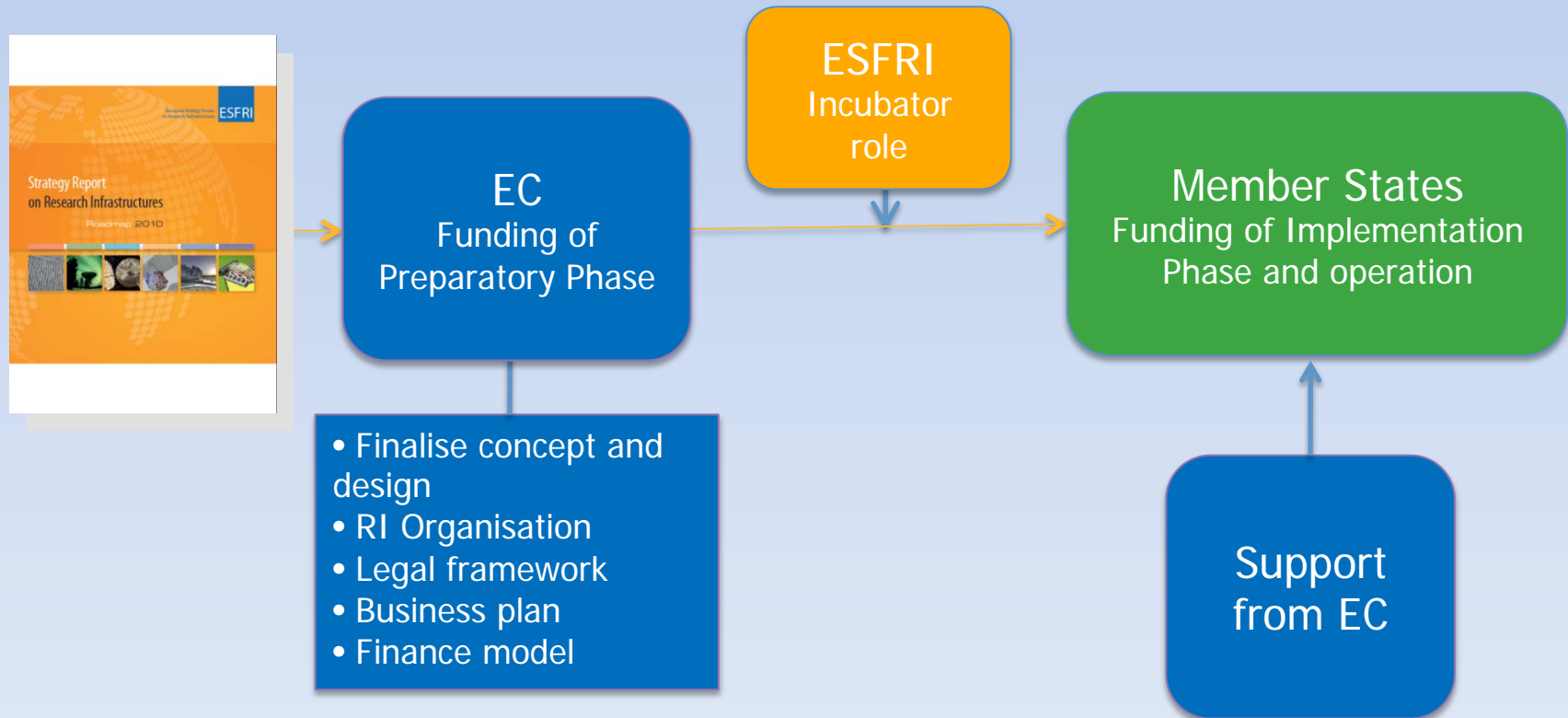- ◦ Evaluation and new Roadmap in the plans (... 2005)

ESFRI – The Roadmap

- New pan-European RIs or major up-grades to existing Ris

- Needs of European research communities in the next 10-20 years

- All fields of Sciences and Technologies, regardless of possible location

- Published: 2006, updated: 2008 & 2010
  - 48 projects
  - Financial investment: ~20 b€, long term commitment for operation: ~2 b€/year

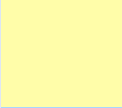European Strategy Forum on Research Infrastructures **ESFRI**

European Strategy Forum on Research Infrastructures **ESFRI**

European Strategy Forum on Research Infrastructures **ESFRI**

Strategy Report on Research Infrastructures
Roadmap 2010

The ESFRI Process (1)

RI proposal by Member State(s) → ESFRI

ESFRI → TWG1, TWG2, TWG3

TWG2 →
- Pan-European
- Science case
- Technical case
- Business case
- Overall maturity

Recommended projects and analysis of landscape

Strategy Report on Research Infrastructures — Roadmap 2010 — ESFRI

The ESFRI Process (2)

EC — Funding of Preparatory Phase

- Finalise concept and design
- RI Organisation
- Legal framework
- Business plan
- Finance model

ESFRI Incubator role

Member States — Funding of Implementation Phase and operation

Support from EC

ESFRI Roadmap 2010

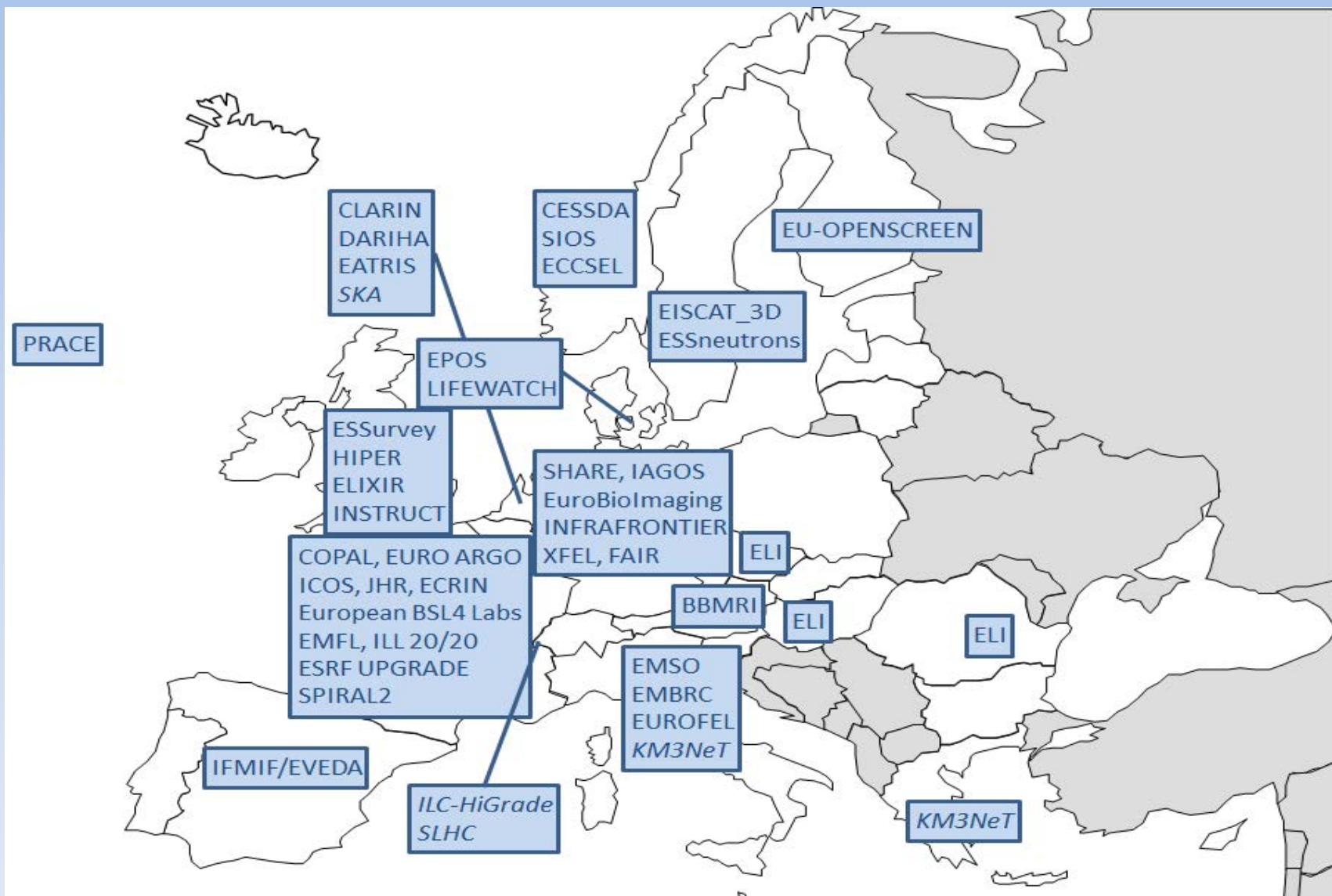| Social Sc. & Hum. (5)	Life Sciences (13)		Environmental Sciences (9)		Energy (7)	Material and Analytical Facilities (6)	Physics and Astronomy (10)		e-Infra-structures (1)
SHARE	BBMRI	ELIXIR	ICOS	EURO-ARGO	ECCSEL	EUROFEL	ELI	TIARA	PRACE
European Social Survey	ECRIN	INFRA FRONTIER	LIFEWATCH	IAGOS	Windscanner	EMFL	SPIRAL2	CTA	
CESSDA	INSTRUCT	EATRIS	EMSO	EPOS	EU-SOLARIS	European XFEL	E-ELT	SKA	
CLARIN	EU-OPENSCREEN	EMBRC	SIAEOS	EISCAT_3D	JHR	ESRF Upgrade	KM3NeT	FAIR	
DARIAH	Euro BioImaging	ERINHA BSL4 Lab		COPAL	IFMIF	NEUTRON ESS	SLHC-PP	ILC-HIGRADE	
	ISBE	MIRRI			HiPER	ILL20/20 Upgrade			
	ANAEE				MYRRHA				

Distributed research infrastructures

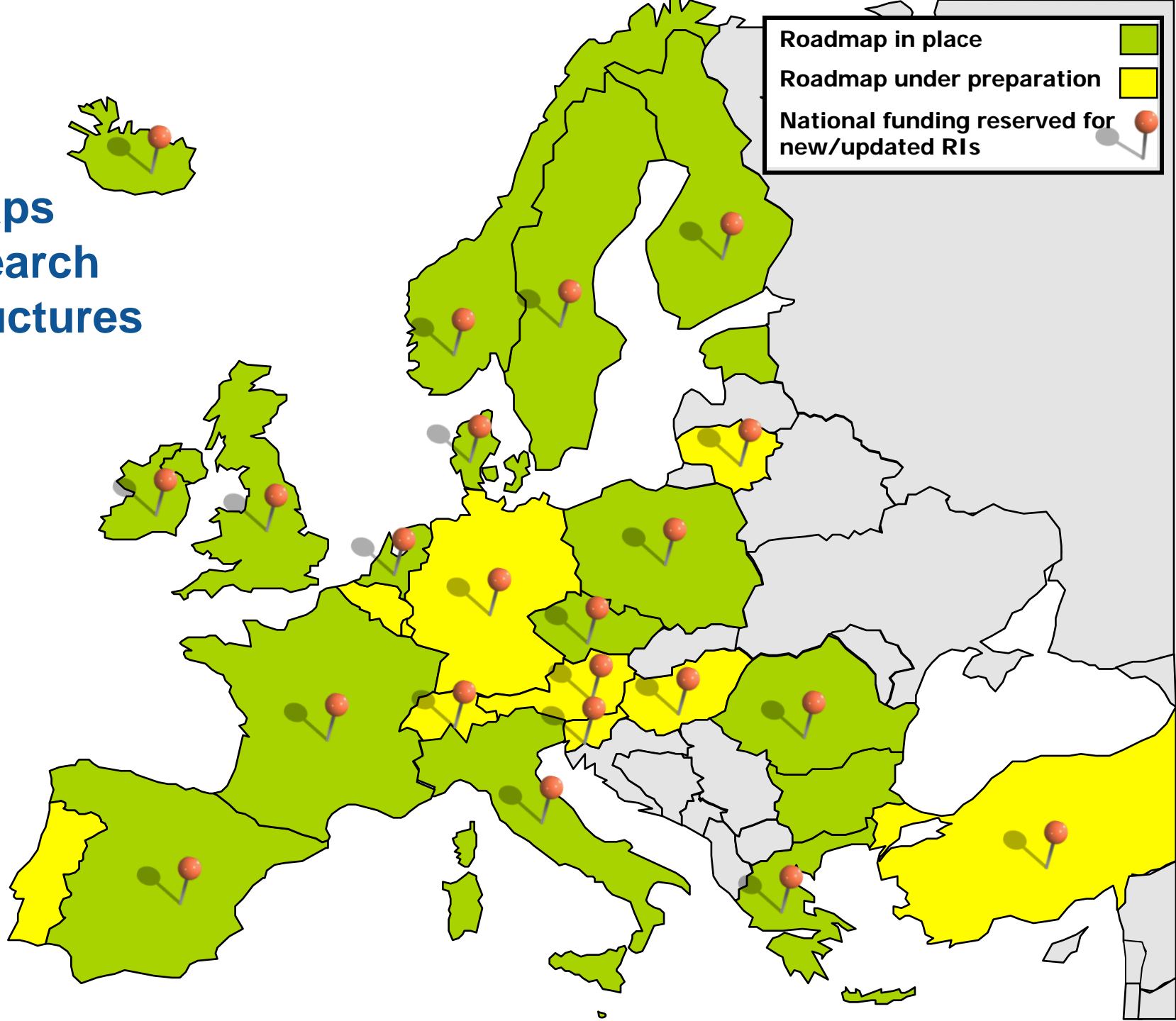Single sited research infrastructures

MAD:IK

Many <u>ARE</u> e-Infras

- CLARIN: Language resources ... enabling eHumanities
- DARIAH: Digital RI for the Arts & Humanities
- LIFEWATCH: An e-science and tech infrastructure for biodiversity and ecosystem research
- ELIXIR: Platform for biological data
- ...
- PRACE: HPC

All need e-Infras

- EURO-ARGO: Ocean floor - not only instruments … but associated data streams and data centers

- IFMIF: Generation of a materials irradiation database … for fusion reactors

- BBMRI: Bio-banks, bio-molecular resource centers … and bio-computational tools

- …

National Roadmaps for Research Infrastructures

Legend:
- Roadmap in place
- Roadmap under preparation
- National funding reserved for new/updated RIs

What is "e-Infrastructure"?

"an environment where research resources (hardware, software and content) can be readily shared and accessed wherever this is necessary to promote better and more effective research".
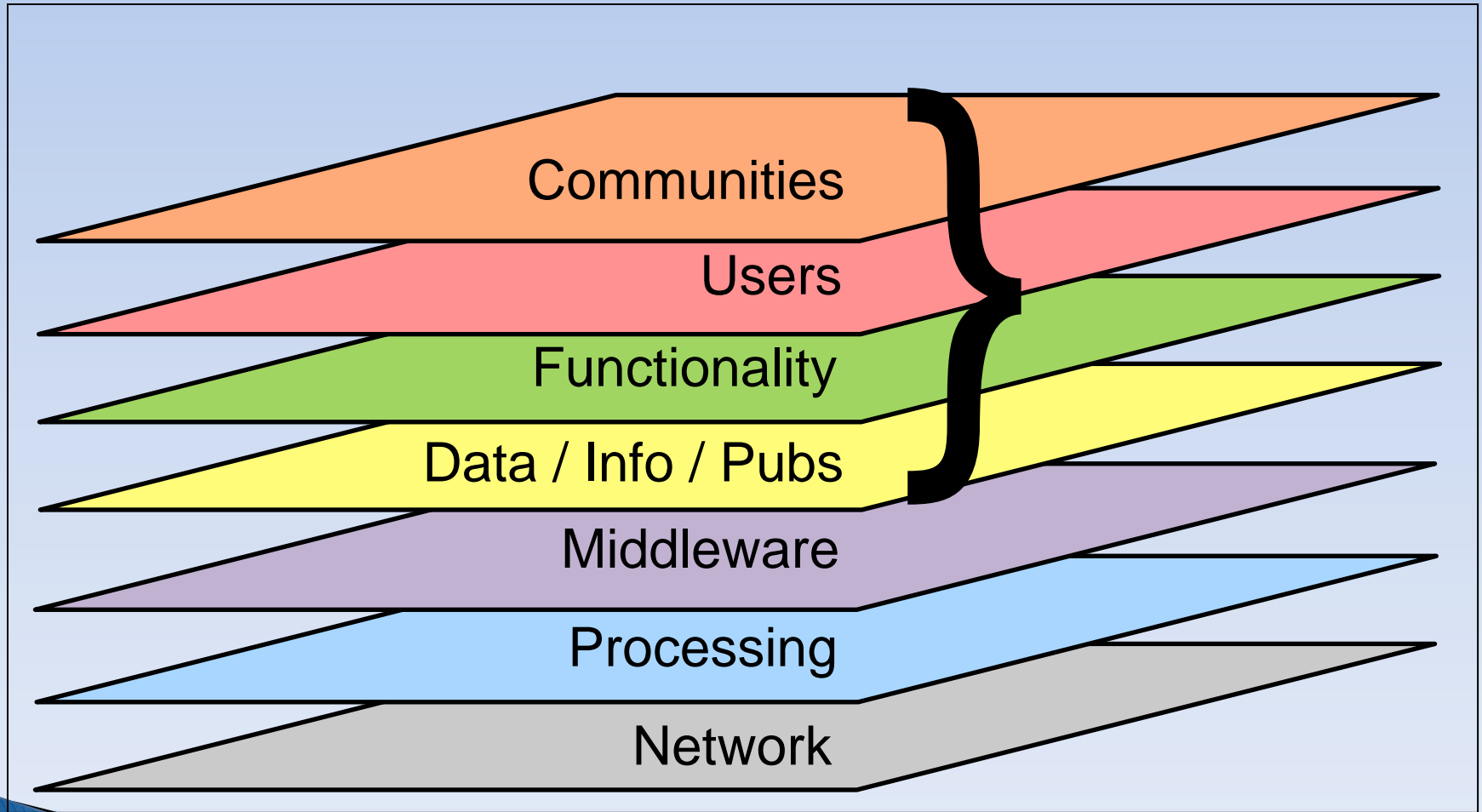
Networks, grids and middleware, comp. resources, exp workbenches, data reps, tools and instruments and the op support that enable global virtual research collabs.

Source: ICT infrastructures for e-Science

eInfrastructure Layers

Common ESFRI Reqs for e-RIs

- Single sign-on: consistent access to resources
- Virtual organisations (collaboration)
- Persistent storage: long-term preservation of data and its access
- Data Management services
- Standards – web services
- Workflows – support of access to HPC/grid/ network resources (compute+data) across Europe
- Training
- Global scope: beyond Europe

Research Paradigms

- Hypothesis ➜ Experiment/Observation ➜ Data collection

- Data collection ➜ Data analysis ➜ Hypothesis

Strategic Decisions Needed

- Technological
- Architectural
- Financial
- Political
- Legal
- Cultural / Social
- ...al

MAD IK

BIG DATA

- Volume (high)
- Velocity (high)
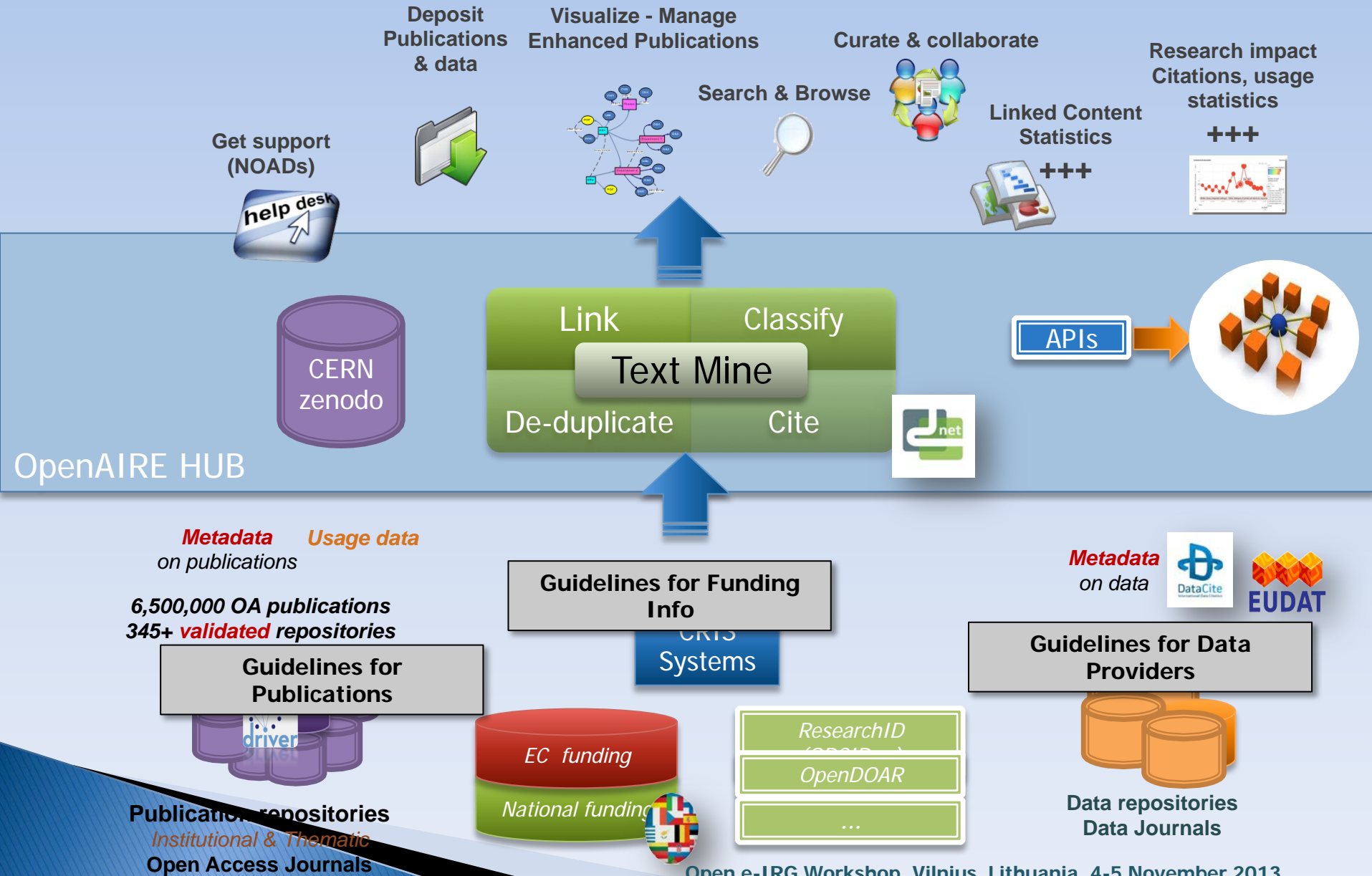- Variety (great)
- Veracity (lack of)
- Value (hard to extract)

Architectural

- Where does data reside?
- Where does computation occur?

- Past/present researcher common behavior
  - discover datasets through metadata queries
  - download *all* data for local processing
  - keep analysis results locally
- Moving data to computation does not scale
- BIG DATA

Architectural

- Future researcher behavior
  - discover datasets through metadata and data queries
  - upload computation to RI for remote processing

  - keep analysis results at RI
  - keep computation at RI
- Computation and data e-infras together
- Cultural/social issues
- Political/legal issues

OpenAIRE

Deposit Publications & data

Visualize - Manage Enhanced Publications

Curate & collaborate

Research impact Citations, usage statistics

Search & Browse

Linked Content Statistics

+++

+++

Get support (NOADs)

help desk

OpenAIRE HUB

CERN zenodo

Link	Classify
Text Mine	
De-duplicate	Cite

net

APIs

Metadata   *Usage data*
on publications

6,500,000 OA publications
345+ validated repositories

Guidelines for Funding Info

CRIS Systems

Metadata
on data

DataCite

EUDAT

Guidelines for Publications

Guidelines for Data Providers

driver

EC funding

National funding

ResearchID

OpenDOAR

…

Publication repositories
Institutional & Thematic
Open Access Journals

Data repositories
Data Journals

Thank you!