e-Infrastructures in Horizon 2020

Vision, approach, drivers, policy background, challenges, WP structure
INFODAY – France
Paris, 25 mars 2014

Jean-Luc Dorel
European Commission – DG CNECT
eInfrastructure

Approach

Transversal
> Cutting across disciplines and sectors

Support tomorrow's science
> Open science, open access, best solutions

Enabling innovation
> Developing and testing innovative solutions
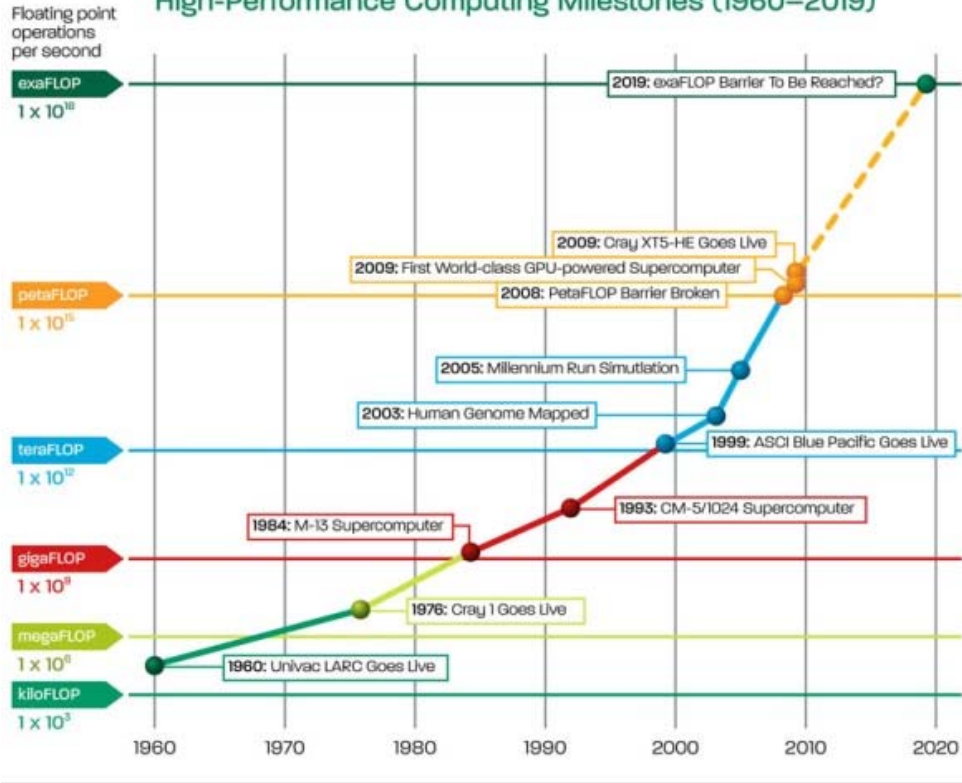> Servicing industry and SMEs
> Spinning out technologies

DRIVERS for change

- COMPUTATIONAL CAPABILITIES

High-Performance Computing Milestones (1960–2019)

Floating point operations per second

exaFLOP 1 x 10^18 — 2019: exaFLOP Barrier To Be Reached?

2009: Cray XT5-HE Goes Live
2009: First World-class GPU-powered Supercomputer
2008: PetaFLOP Barrier Broken

petaFLOP 1 x 10^15

2005: Millennium Run Simulation
2003: Human Genome Mapped
1999: ASCI Blue Pacific Goes Live

teraFLOP 1 x 10^12

1993: CM-5/1024 Supercomputer
1984: M-13 Supercomputer

gigaFLOP 1 x 10^9

1976: Cray 1 Goes Live

megaFLOP 1 x 10^6

1960: Univac LARC Goes Live
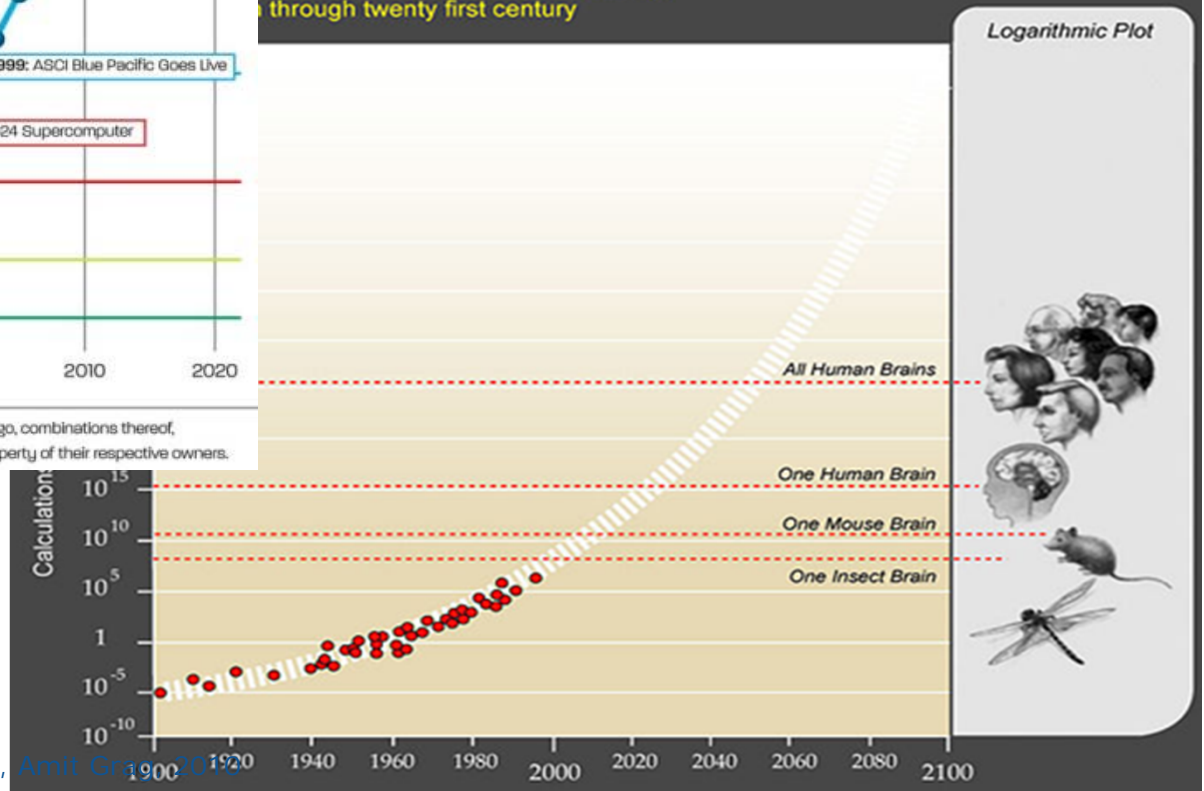
kiloFLOP 1 x 10^3

1960 1970 1980 1990 2000 2010 2020

©2010 Advanced Micro Devices, Inc. All rights reserved. AMD, the AMD Arrow logo, combinations thereof, are trademarks of Advanced Micro Devices, Inc. All other trademarks are the property of their respective owners.

MIT Technology Review, 2012

MORE COMPUTING POWER

...ential Growth of Computing
... through twenty first century

Logarithmic Plot

All Human Brains
One Human Brain
One Mouse Brain
One Insect Brain

10^{15}
10^{10}
10^{5}
1
10^{-5}
10^{-10}

Calculations

1900 1920 1940 1960 1980 2000 2020 2040 2060 2080 2100

Future of learning technology - 2015, Amii Grego...

DRIVERS for change

- **COMPUTATIONAL CAPABILITIES**
- **BIG DATA**

BIG DATA ...

GROWTH OF THE DIGITAL UNIVERSE, 2010-2020

Digital Universe in Exabytes (Billions of Gigabytes)

2005
130

2010
1,227

2012
2,837

2015
8,591

2020
40,026

Business and the Digital Universe, IDC, 2012

DRIVERS for change

- COMPUTATIONAL CAPABILITIES
- BIG DATA
- GLOBAL CONNECTIONS
- GLOBAL PARTICIPATION

GLOBAL CONNECTIONS ...

Map of scientific collaborations from 2005 to 2009
Computed by Olivier H. Beauchesne @ Science-Metrix, Inc.
Data from Scopus, using books, trade journals and peer-reviewed journals

Horizon 2020: What's new?

A single programme bringing together three separate programmes/initiatives*

More innovation, from research to retail, all forms of innovation

Focus on societal challenges facing EU society, e.g. health, clean energy and transport

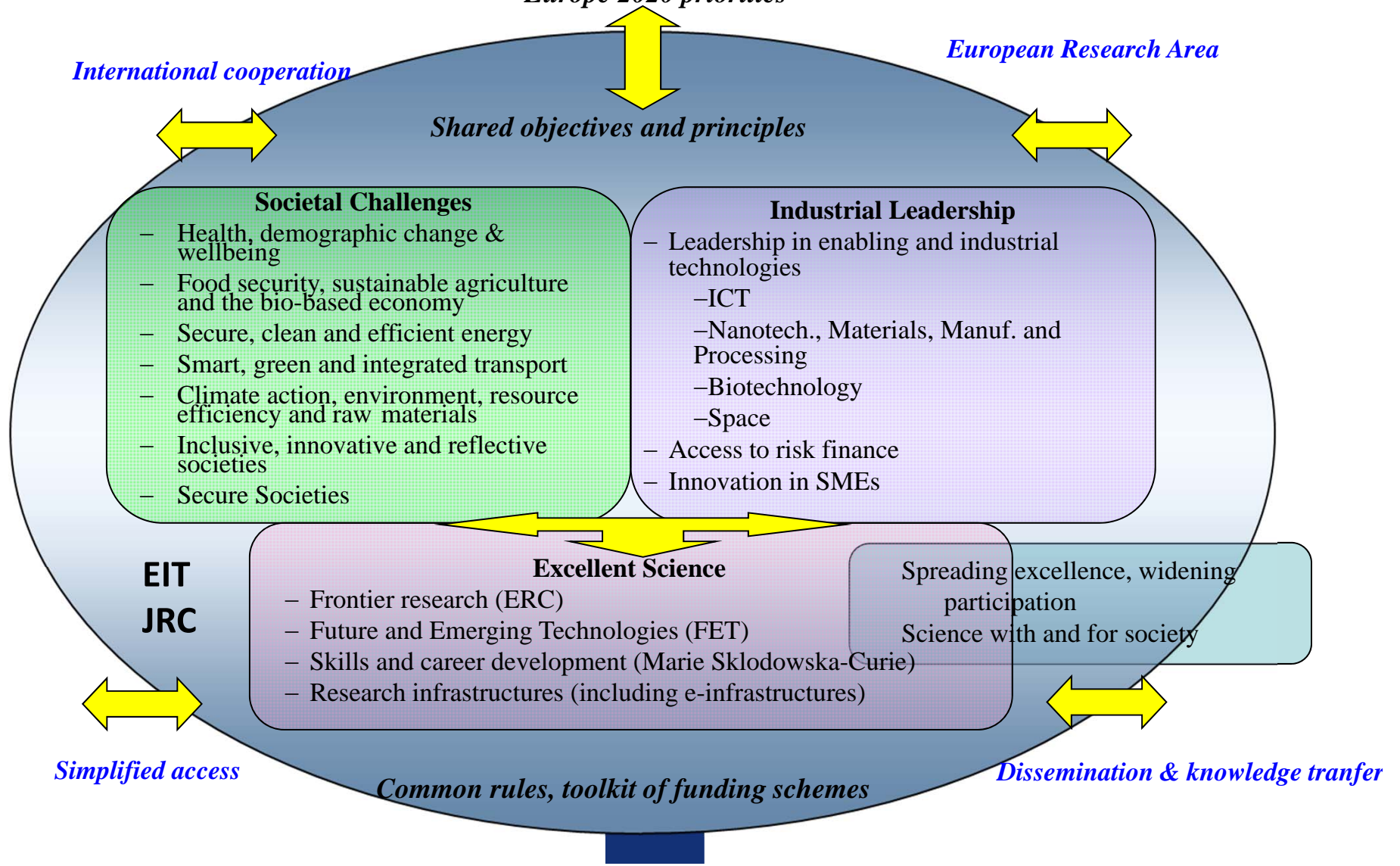*Simplified access,* for all companies, universities, institutes in all EU countries and beyond.

The 7th research Framework Programme (FP7), innovation aspects of Competitiveness and Innovation Framework Programme (CIP), EU contribution to the European Institute of Innovation and Technology (EIT)
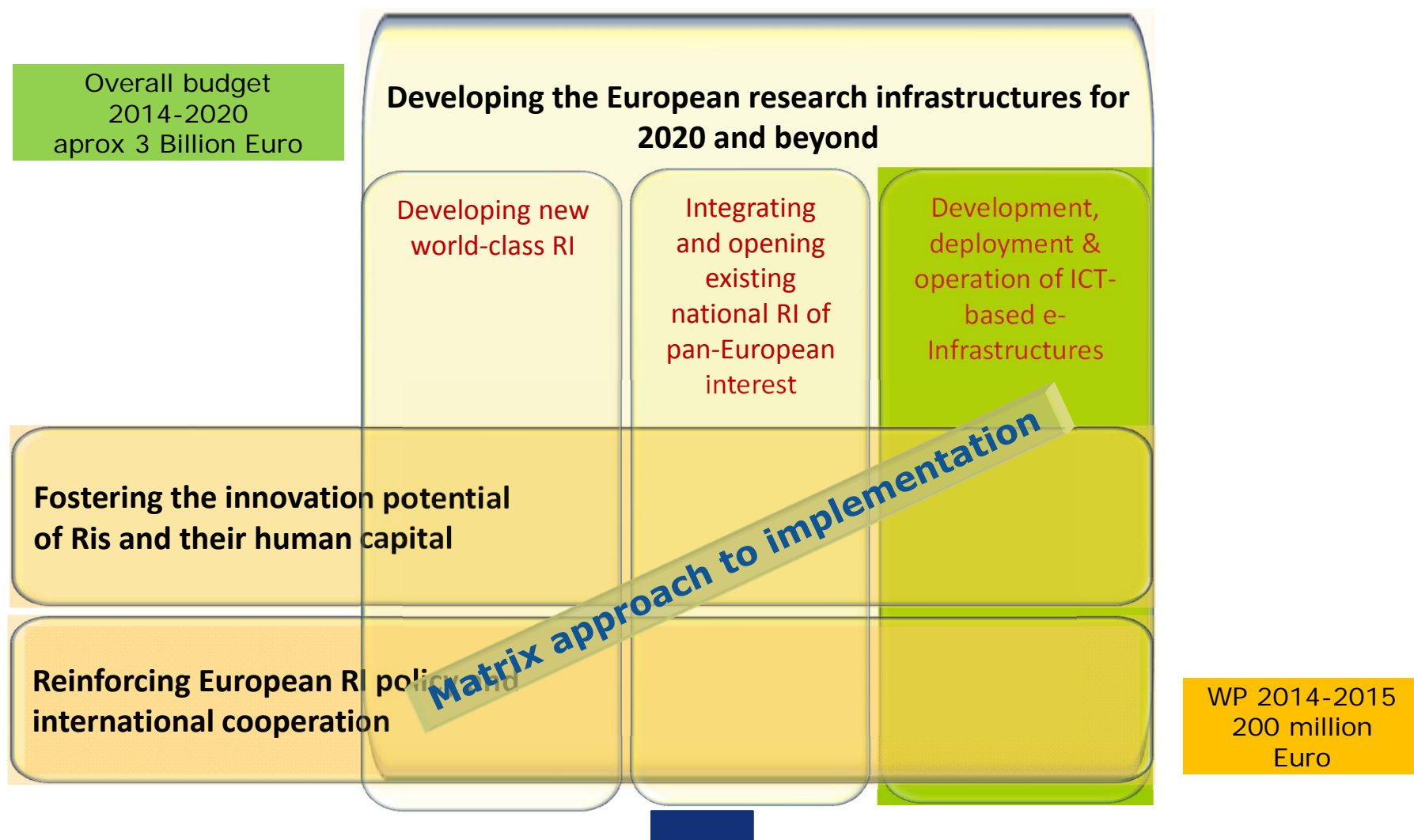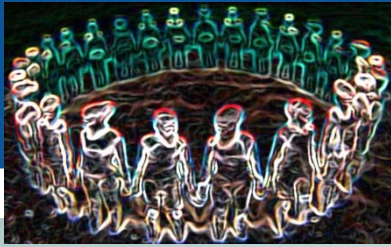
European Commission

Europe 2020 priorities

International cooperation

European Research Area

Shared objectives and principles

Societal Challenges
- Health, demographic change & wellbeing
- Food security, sustainable agriculture and the bio-based economy
- Secure, clean and efficient energy
- Smart, green and integrated transport
- Climate action, environment, resource efficiency and raw materials
- Inclusive, innovative and reflective societies
- Secure Societies

Industrial Leadership
- Leadership in enabling and industrial technologies
  - ICT
  - Nanotech., Materials, Manuf. and Processing
  - Biotechnology
  - Space
- Access to risk finance
- Innovation in SMEs

EIT
JRC

Excellent Science
- Frontier research (ERC)
- Future and Emerging Technologies (FET)
- Skills and career development (Marie Sklodowska-Curie)
- Research infrastructures (including e-infrastructures)

Spreading excellence, widening participation
Science with and for society

Simplified access

Common rules, toolkit of funding schemes

Dissemination & knowledge tranfer

Research Infrastructures in Horizon 2020

Overall budget 2014-2020 aprox 3 Billion Euro

Developing the European research infrastructures for 2020 and beyond

Developing new world-class RI

Integrating and opening existing national RI of pan-European interest

Development, deployment & operation of ICT-based e-Infrastructures

Fostering the innovation potential of Ris and their human capital

Reinforcing European RI policy and international cooperation

Matrix approach to implementation

WP 2014-2015 200 million Euro

Policy Background (1/3):
Research Data become an infrastructure for modern science

Europe is "Riding the Wave" Report

- Data e-infrastructure that supports seamless access, use, re-use and trust of data

- Physical and technical infrastructure become invisible and the **data becomes the infrastructure**

Commission Communication on Scientific Information COM(2012)401

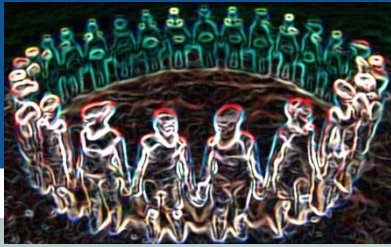- Access, preservation and e-infrastructure (publications and data)

ERA Communication COM(2012)392

- Federation of researcher electronic identities

Riding the wave
How Europe can gain from the rising tide of scientific data

Final report of the High Level Expert Group on Scientific Data
A submission to the European Commission
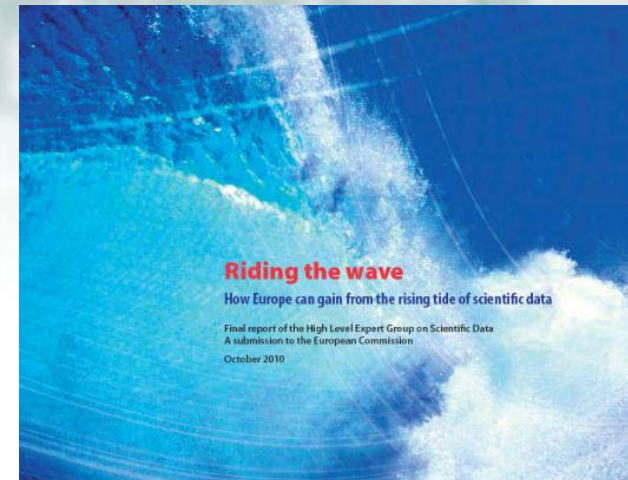
October 2010

Riding the Wave
High Level Expert Group on Scientific
Data, October 2010

http://cordis.europa.eu/fp7/ict/e-infrastructure/docs/hlg-sdi-report.pdf

Policy Background 2:
European HPC Strategy –integrated approach in H2020

Basis: Commission Communication "High-Performance Computing: Europe's place in a Global Race" (2012)

Vision: to ensure European leadership in the <u>supply</u> and <u>use</u> of HPC systems and services by 2020 in a strategy combining:

(a) developing the next generation of HPC towards exascale;

(b) providing access to the best HPC infrastructure for both industry and academia;

(c) achieving excellence in computing applications - existing or new – driven by the needs of science, industry and SMEs

Linking demand and supply – in the spirit of Horizon 2020

Contractual Public-Private Partnership (cPPP) covering (a) and part of (c)

HPC - Examples of interrelations between actions

PRACE

Access to best HPC for industry and academia

ETP4HPC/cPPP

Autonomous EU development of Exascale technologies

- specifications of exascale prototypes forTier-0
- technological options for future procurements

- CoEs may be associated to PRACE Centres
- provision of HPC capabilities and expertise

- identify applications for co-design of exascale systems

Centres of Excellence

Network of SME Competence Centres

Excellence in HPC application

Policy Background (3/3): GÉANT Expert Group Report

World Class Connectivity and Services to Knowledge Communities

- Support *Growth and Opening up*
- Help to close *digital divides*
- Europe as global *hub*
- Stimulate *innovation*

Reorganize for 2020

- *Flexibility* in Technology and Architecture
- *Experimentation* and *standardisation*
- Improve *Governance*
- Step up *Funding*
- Update the *Regulatory* Regime

Knowledge without Borders

GÉANT 2020 as the
European Communications Commons

Report of the GÉANT Expert Group
••• October 2011

European Commission
Information Society and Media

INTEGRATED e-INFRASTRUCTURE SERVICES

VRE

E-INFRASTRUCTURES FOR VIRTUAL RESEARCH ENVIRONMENTS (VRE)

PROVISION OF CORE SERVICES ACROSS E-INFRASTRUCTURES

DATA

COMPUTING

NETWORK OF HPC COMPETENCE CENTRES FOR SMES

MANAGING, PRESERVING AND COMPUTING WITH BIG RESERACH DATA

RESEARCH AND EDUCATION NETWORKING – GEANT

CENTRES OF EXCELLENCE FOR COMPUTING APPLICATIONS

E-INFRASTRUCTURES FOR OPEN ACCESS

CONNECTIVITY

PAN-EUROPEAN HIGH PERFORMANCE COMPUTING INFRASTRUCTURE AND SERVICES

TOWARDS GLOBAL DATA E-INFRASTRUCTURES RESEARCH DATA ALLIANCE

SUPPORT

E-INFRASTRUCTURE POLICY DEVELOPMENT AND INTERNATIONAL COOPERATION

NEW PROFESSIONS AND SKILLS FOR E-INFRASTRUCTURES

European Commission
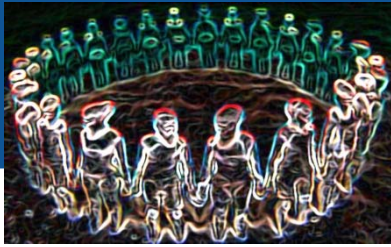
RESEARCH INFRASTRUCTURE
Work Programme 2014-2015

CALL 3 → E-INFRASTRUCTURES

MANAGING, PRESERVING AND COMPUTING WITH BIG RESERACH DATA	E-INFRASTRUCTURES FOR OPEN ACCESS	TOWARDS GLOBAL DATA E-INFRASTRUCTURES: RESEARCH DATA ALLIANCE	PAN-EUROPEAN HIGH PERFORMANCE COMPUTING INFRASTRUCTURE AND SERVICES

| CENTRES OF EXCELLENCE FOR COMPUTING APPLICATIONS	NETWORK OF HPC COMPETENCE CENTRES FOR SMES	PROVISION OF CORE SERVICES ACROSS E-INFRASTRUCTURES	RESEARCH AND EDUCATION NETWORKING – GEANT	E-INFRASTRUCTURES FOR VIRTUAL RESEARCH ENVIRONMENTS (VRE)

CALL 4 → SUPPORT TO INNOVATION, HUMAN RESOURCES, POLICY AND INTERNATIONAL COOPERATION FOR RESEARCH INFRASTRUCTURES

E-INFRASTRUCTURE POLICY DEVELOPMENT AND INTERNATIONAL COOPERATION	NEW PROFESSIONS AND SKILLS FOR E-INFRASTRUCTURES
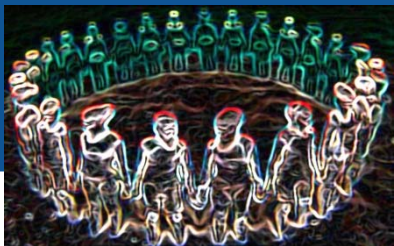
INTEGRATED
e-INFRASTRUCTURE
SERVICES

across...

Industry sectors   disciplines

societal chalenges

European Commission

Automotive

Oil and gas

Pharmaceuticals

Smart cities

Health

Climate change

Physics

Social sciences

... ...

SUMMARY ...

Open access in Horizon 2020 and in the ERA

European Commission

What is Open Access (OA)?

OA = online access at no charge to the user

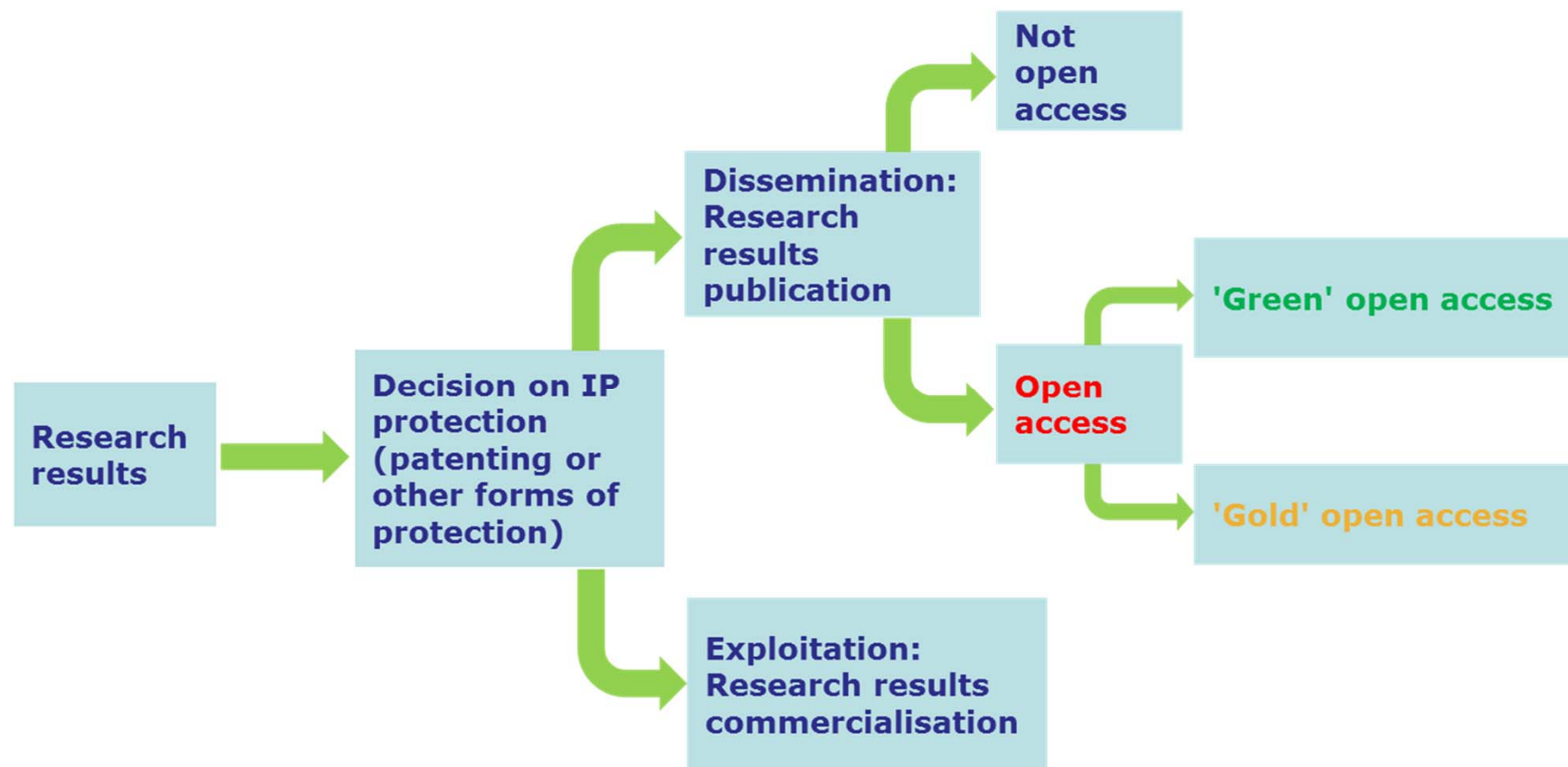- to peer-reviewed scientific publications
- to research data

Two main OA publishing business models

- **Self-archiving**: deposit of manuscripts & **immediate/delayed OA** provided by author ("Green OA")
- **OA publishing**: costs covered & **immediate OA** provided by publisher ("Gold OA")

What OA is NOT

- **Not an obligation to publish**
- **Not at odds with patenting**
- **OA publications go the same peer review process**

Dissemination and exploitation

OA to publications: H2020 mandate (1)

Each beneficiary must ensure OA to all peer-reviewed scientific publications relating to its results:

- **Deposit** a machine-readable copy of the published version or final peer-reviewed manuscript accepted for publication in a repository of the researchers choice (possibly OpenAIRE compliant)

- **Ensure OA** on publication or at the latest within 6 months (12 for SSH)

- **Aim to deposit** at the same time **the research data needed to validate the results ("underlying data")**

- Ensure **OA to the bibliographic metadata** that identify the deposited publication, via the repository

OA to publications: H2020 mandate (2)

Routes towards OA:

- OA publishing and self-archiving considered valid and complementary routes

- Deposit into a repository also in the case of OA publishing

Costs for OA publishing:

- Eligibility of OA publishing costs during the grant (as in FP7)

- Piloting a mechanism for open access publishing after the end of the grant agreement (call EINFRA-2-2014 – eInfrastructure for Open Access)

Licencing:

- Encouragement to authors to retain their copyright and grant adequate licences to publishers (e.g. Creative Commons)

Pilot on Open Research Data (1)

Areas of the 2014-2015 Work Programme participating in the Open Research Data Pilot are:

- Future and Emerging Technologies

- Research infrastructures – part e-Infrastructures

- Leadership in enabling and industrial technologies – Information and Communication Technologies

- Societal Challenge: Secure, Clean and Efficient Energy – part Smart cities and communities

- Societal Challenge: Climate Action, Environment, Resource Efficiency and Raw materials – except raw materials

- Societal Challenge: Europe in a changing world – inclusive, innovative and reflective Societies

- Science with and for Society

Projects in other areas can participate on a voluntary basis.

Pilot on Open Research Data (2)

Types of data concerned:

- Data (including associated metadata) needed to validate the results presented in scientific publications ("underlying data")

- Other data (including associated metadata) as specified in data management plan

Beneficiaries participating in the Pilot will:

- Deposit this data in a research data repository of their choice (e.g. openaire)

- Take measures to make it possible to access, mine, exploit, reproduce and disseminate free of charge (using e.g. Creative Commons licences)

- Provide information about tools and instruments at the disposal of the beneficiaries and necessary for validating the results (where possible, provide the tools and instruments themselves)

Support & monitoring to be developed

Pilot on Open Research Data (3)

Projects may opt out of the Pilot on Open Research Data in Horizon 2020 in a series of cases:

- If the project will not generate / collect any data

- Conflict with obligation to protect results

- Conflict with confidentiality obligations

- Conflict with security obligations

- Conflict with rules on protection of personal data

- If the achievement of the action's main objective would be jeopardised by making specific parts of the research data openly accessible (to be explained in data management plan)

Data management in Horizon 2020

- **Data Management Plans (DMPs) mandatory for all projects participating in the pilot (deliverable within the first six months)**

- **Other projects invited to submit a DMP if relevant for their planned research**

- DMP questions:

    - What data will be collected / generated?

    - What standards will be used / how will metadata be generated?

    - What data will be exploited? What data will be shared/made open?

    - How will data be curated and preserved?

Proposal ID	Acronym	Go to

5 - Call specific questions

Open Research Data Pilot in Horizon 2020

If selected, all applicants will participate in the Pilot on Open Research Data in Horizon 2020[1], which aims to improve and maximise access to and re-use of research data generated by actions. Participating in the Pilot does not necessarily mean opening up all research data. Actions participating in the Pilot will be invited to formulate a Data Management Plan in which they will determine and explain which of the research data they generate will be made open.

Applicants have the possibility to opt out of this Pilot and must indicate a reason for this choice.

Participation in this Pilot does not constitute part of the evaluation process. Proposals will not be evaluated favourably because they are part of the Pilot and will not be penalised for opting out of the Pilot.

We wish to opt out of the Pilot on Open Research Data in Horizon 2020.	● Yes	○ No

Please indicate the reason(s) for not being able to participate in the Pilot:

- the project does not generate any data	☐
- to allow the protection of results (e.g. patenting)	☐
- incompatibility with the need for confidentiality linked to security reasons	☐
- incompatibility with privacy/data protection reasons	☐
- achievement of the project's main aim would be jeopardised	☐
- other legitimate reasons	☒

Please specify the reason:

> *Other legitimate reasons (Maximum 300 characters)*

Remaining characters   300

[1] According to article 43.2 of Regulation (EU) No 1290/2013 of the European Parliament and of the Council, of 11 December 2013, laying down the rules for participation and dissemination in

---

Proposal ID	Acronym	Go to

Data management activities

The use of a Data Management Plan (DMP) is required for projects participating in the Open Research Data Pilot in Horizon 2020, in the form of a deliverable in the first 6 months of the project.

All other projects may deliver a DMP on a voluntary basis, if relevant for their research.

Are data management activities relevant for your proposed project?	● Yes	○ No

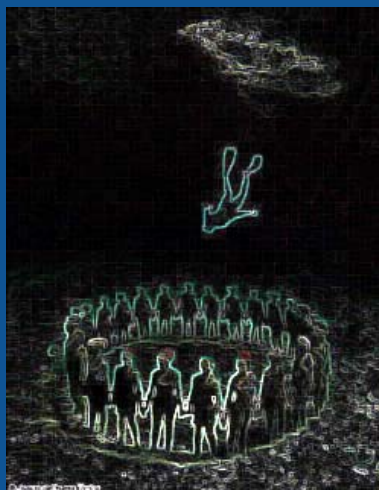| A Data Management Plan will be delivered (Please note: Projects participating in the Open Research Data Pilot **must** include a Data Management Plan as a deliverable in the first 6 months of the project).	☐
Data Management is part of a Work Package.	☐
Data Management will be integrated in another way.	☒

Please indicate how data management will be addressed in your project:

> *Please indicate how data management will be addressed in your project.*

Remaining characters   300

Proposal ID	Acronym	Go to

5 - Call specific questions

Open Research Data Pilot in Horizon 2020

If selected, all applicants have the possibility to participate in the Pilot on Open Research Data in Horizon 2020[1], which aims to improve and maximise access to and re-use of research data generated by actions. Participating in the Pilot does not necessarily mean opening up all research data. Actions participating in the Pilot will be invited to formulate a Data Management Plan in which they will determine and explain which of the research data they generate will be made open.

We wish to participate in the Pilot on Open Research Data in Horizon 2020 on a voluntary basis	● Yes	○ No

Participation in this Pilot does not constitute part of the evaluation process. Proposals will not be evaluated favourably because they are part of the Pilot and will not be penalised for not participating.

[1] According to article 43.2 of Regulation (EU) No 1290/2013 of the European Parliament and of the Council, of 11 December 2013, laying down the rules for participation and dissemination in "Horizon 2020 – the Framework Programme for Research and Innovation (2014-2020)" and repealing Regulation (EC) No 1906/2006.

Data management activities

The use of a Data Management Plan (DMP) is required for projects participating in the Open Research Data Pilot in Horizon 2020, in the form of a deliverable in the first 6 months of the project.

All other projects may deliver a DMP on a voluntary basis, if relevant for their research.

Are data management activities relevant for your proposed project?	● Yes	○ No

A Data Management Plan will be delivered (Please note: Projects participating in the Open Research Data Pilot **must** include a Data Management Plan as a deliverable in the first 6 months of the project).	☐
Data Management is part of a Work Package.	☐
Data Management will be integrated in another way.	☒

Please indicate how data management will be addressed in your project:
Please indicate how data management will be addressed in your project.

Remaining characters   300

**Merci de votre attention,
Jean-luc.dorel@ec.europa.eu**

European Commission – DG CNECT
eInfrastructures