



data: an infrastructure for science

Données de la Recherche : enjeux, perspectives, politique(s)
Paris, 29 May 2012

Carlos Morais Pires
European Commission
carlos.morais-pires [at] ec.europa.eu

Digital Agenda
1 0 1 0 1 1 1 0 1 1 1 0 0 0 0 1 0 0 2010-2020
for Europe



Neelie Kroes

Digital Agenda

Digital (information) single market

Open Science means optimal sharing of research results and tools such as publications, research data, software, educational resources and infrastructures across institutional, disciplinary and national boundaries.



Open Science

Open Scientific Content (*data, computational resources and software resulting from public funded research should be made openly available and preserved, for re-use in research and education activities*)

Open Culture (*career systems should support and reward those who participate in the culture of sharing. Open science should inspire the young and enable adequate education to benefit from the abundance of technical tools and scientific information*)

Open Infrastructures (*reliable, high-performance and economically efficient infrastructures*)



the rising tide of data

"A fundamental characteristic of our age is the rising tide of data – global, diverse, valuable and complex. In the realm of science, this is both an opportunity and a challenge."

Riding the Wave report, High-Level Group on Data



scientific information “continuums”

experimental data and publications (new paradigm)

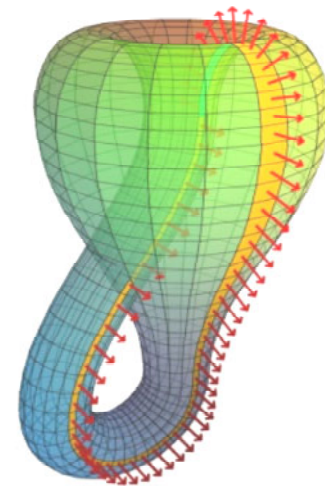
humans and computers (e-infrastructure)

different scientific disciplines (access, multidisciplinary)

past, present and future (preservation)

research and education (public mission)

different institutions (organisation)



Klein Bottle
<http://plus.maths.org/issue26/index.html>

story of the Girl and the Wave

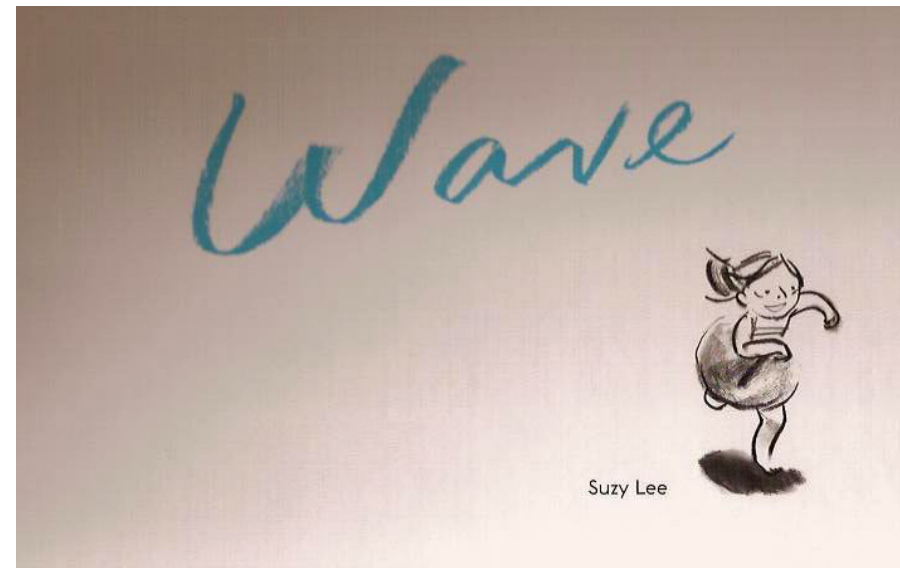
Ocean of Unknown

adapted from the book "Wave" by Suzy Lee

What science knows is just a drop of an immense Ocean.

In the information Age, bits and bits and more bits of data captured by advanced instruments or collected by skilled researchers are an Ocean of unknowns.

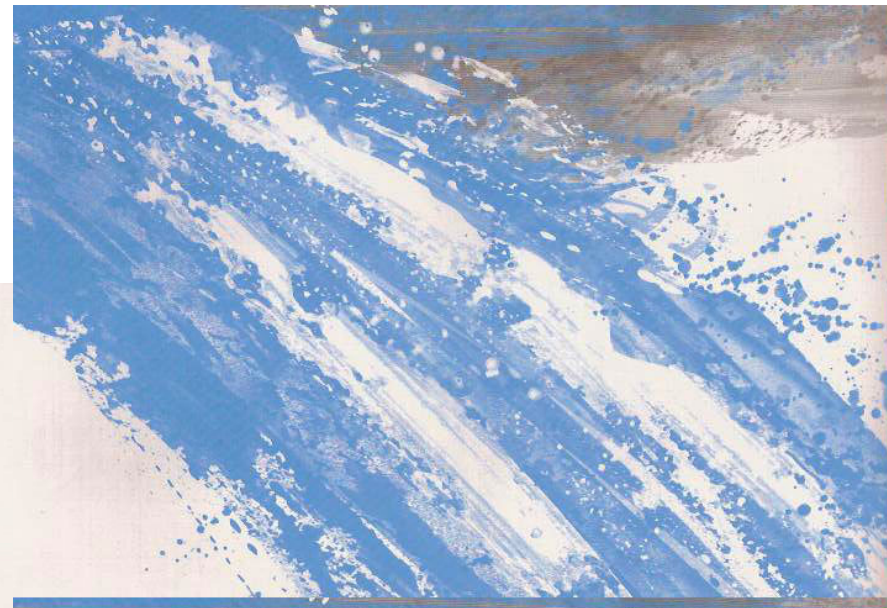
Scientists are trying to unveil what is hidden in that ocean under the small and big waves...



The girl...

*is a young researcher and
dares to challenge big waves*

Challenging the Wave...



Deluge...

The wave hits hard...

The girl was not prepared to deal with such a powerful wave.

But researchers are, by definition, intelligent and they never give up.



New facts...

Serendipity...

The spirit of inquiry, the curiosity, the desire to understand puts back the smile in the girl's face...

And she starts looking into so many things brought by the wave...

The discovery process is in motion...



And finally the rewards...

A culture of sharing...

She dares and she shares...

She is proud to show her data...

She shares it with friends as important evidence of her discoveries...

A smart gull witnesses all and reassures the good quality of the collections.

The world is ready to recognise her merits and her audacity to challenge the waves.



e-Infrastructures for Data

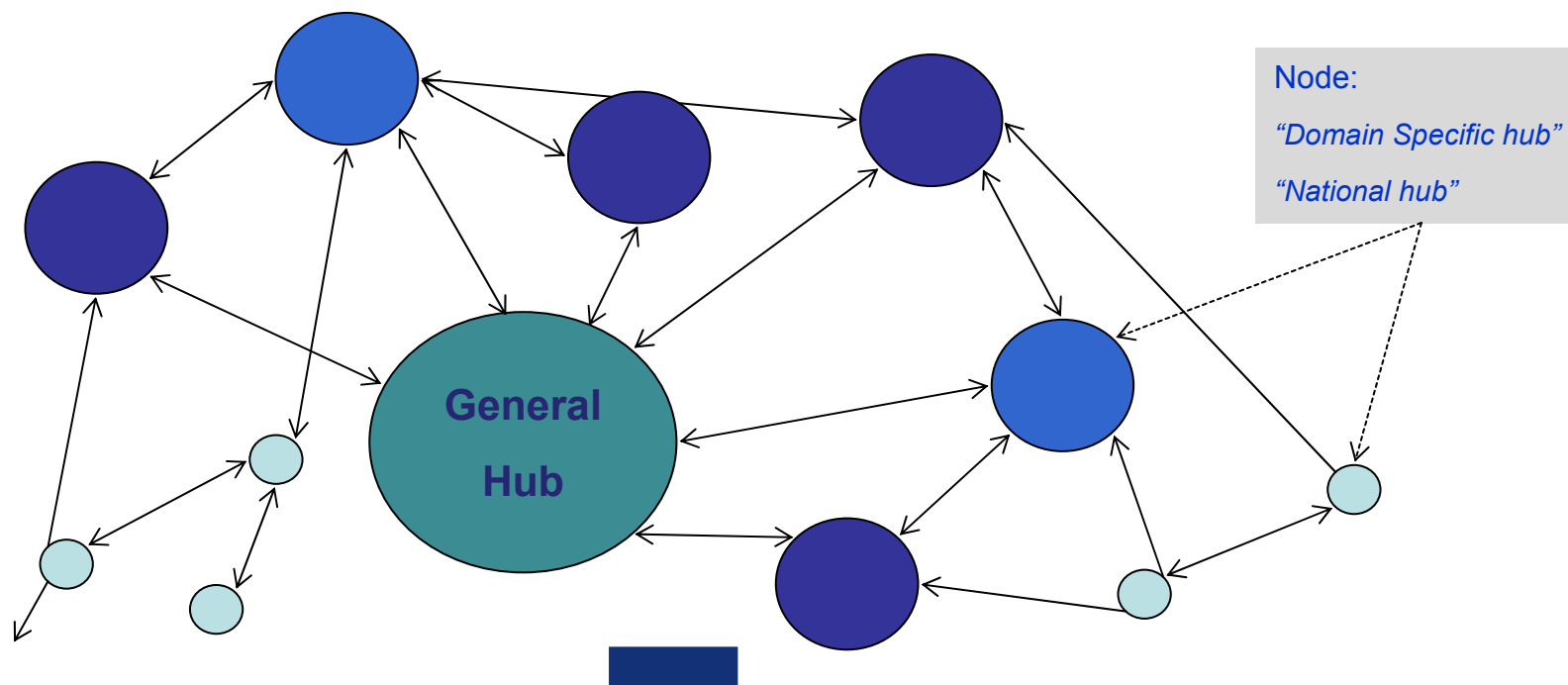
(adapted from Prof. Sulston Presentation in the European Parliament on October 2011)

Distributed and participatory architectures; robust networks of people and institutions



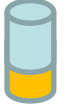


Discoverability, Access and Interoperability of Data

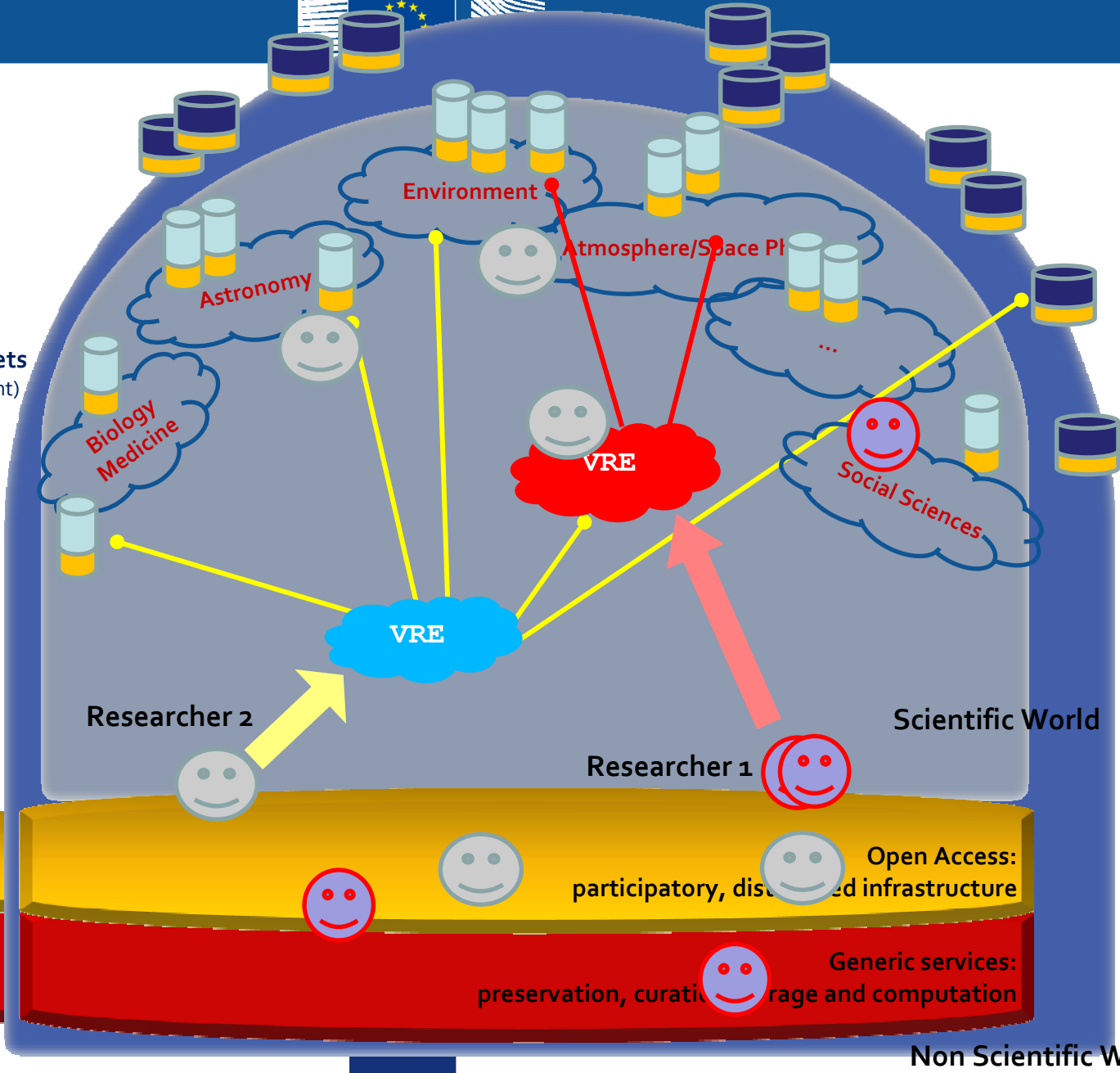
Access to Storage and Computing Resources

High-speed Connectivity to enable international collaborations

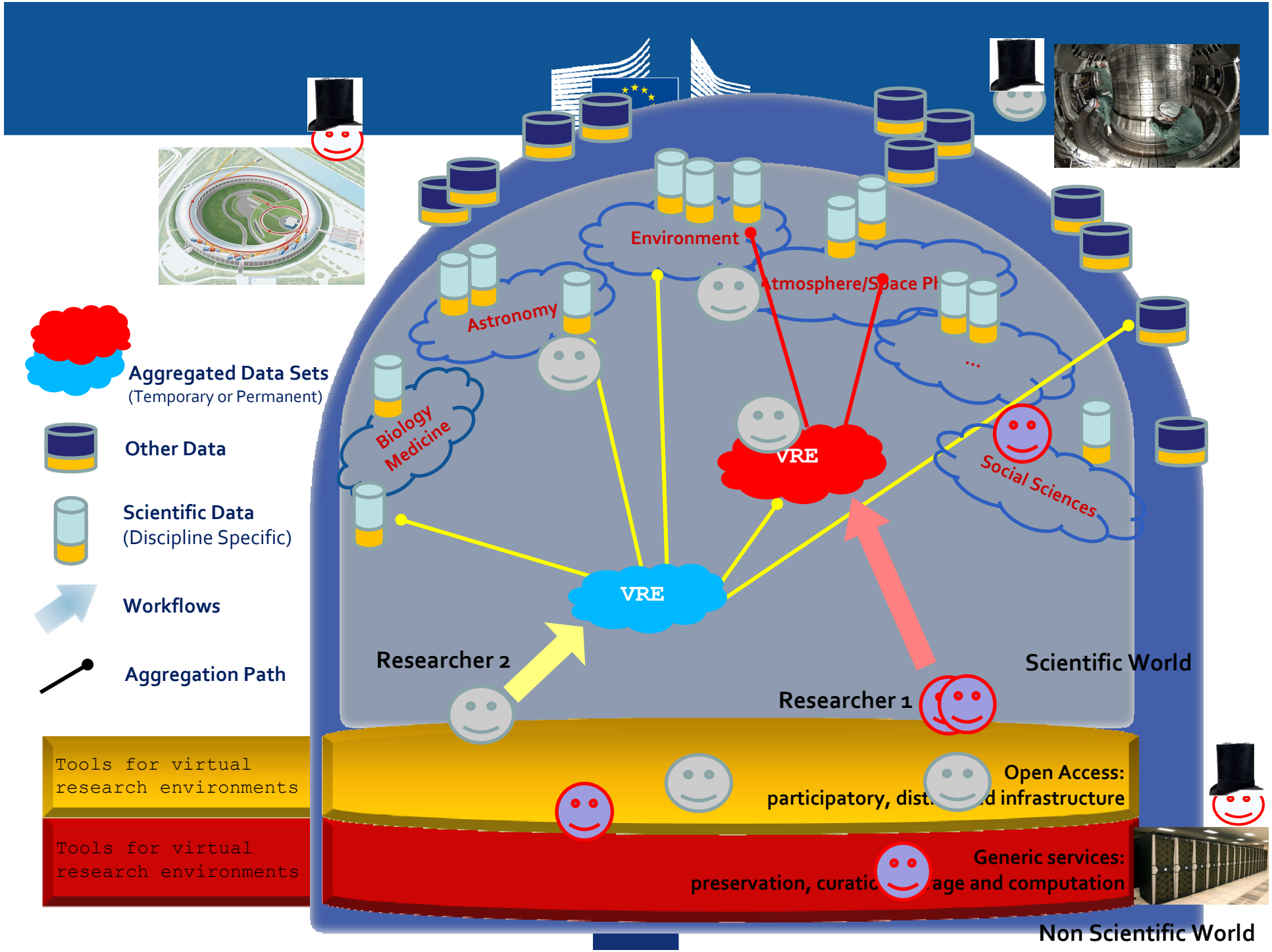




-  **Aggregated Data Sets**
(Temporary or Permanent)
-  **Other Data**
-  **Scientific Data**
(Discipline Specific)
-  **Workflows**
-  **Aggregation Path**



Non Scientific World



Aggregated Data Sets
(Temporary or Permanent)

Other Data

Scientific Data
(Discipline Specific)

Workflows

Aggregation Path

Tools for virtual
research environments

Tools for virtual
research environments

Open Access:
participatory, distributed infrastructure

Generic services:
preservation, curatorial management and computation

Non Scientific World

Researcher 2

Researcher 1

Scientific World

VRE

VRE

Environment

Atmosphere/Space Ph...

Astronomy

Biology
Medicine

Social Sciences



e-Infrastructure for Data



“Our vision is a scientific e-infrastructure that supports seamless access, use, re-use of data.”

Riding the Wave report, High-Level Group on Data

- ***Grand challenges of the 21st century transcend borders***
- ***21st Century science is made of data factories***
- ***Basis for multi disciplinary research***
- ***Global collaboration on Data Access and Interoperability***
- ***Strong commitment to open science***



Scientific Data Infrastructure... ... under construction!



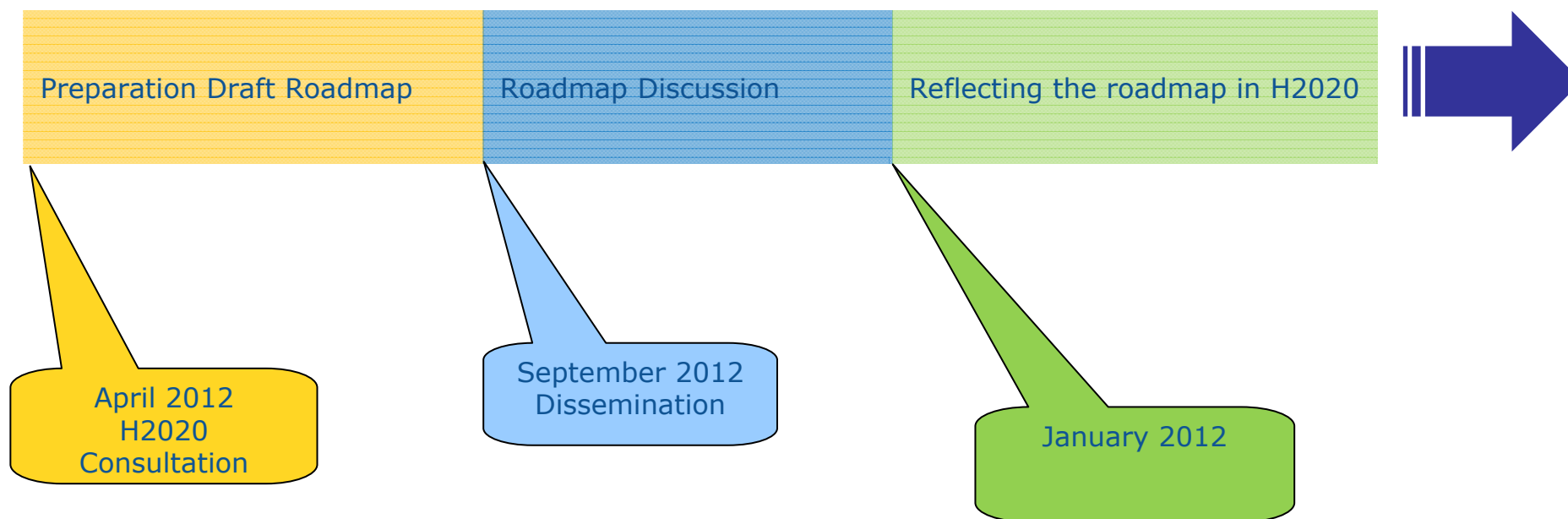
- *Launch European and Global DMP*
- *Enlarge and consolidate e-infrastructure for open access*
- *Set an educational framework for data scientists*
- *Establish pan-European "scholar passport"*
- *Launch community driven e-infrastructure initiatives*
 - *scientific contributions to societal challenges*
- *Launch service-driven e-infrastructure initiatives*
 - *building the capacity for exascale science*
- *Set global coordination mechanism (Data Web Forum)*
- *Support new ways (metrics) to evaluate research results*
- *Take advantage of mobility of researchers/e-infrastructure operators*



need for coordination at European level

e-Infrastructure <i>of</i> Data	Governance (rules for access and preservation)
	Information (Human and Machine)
	Services Manag. of Databases/Repository
e-Infrastructure <i>for</i> Data	Discoverability/Provenance (Metadata, DOIs, DAIs, ...)
	Processing, Computation
	Connectivity/Storage infrastructure

Next Steps



thank you



High-Level Group on Data





Carlos Morais Pires
European Commission
Tel: +32 2 29 63 401

[carlos.morais-pires\(at\)ec.europa.eu](mailto:carlos.morais-pires(at)ec.europa.eu)

Acknowledgements:

The author would like to thank Suzy Lee and the editor suzyleebooks for the permission to use illustrations of Suzy's book "Wave".

