



RESEARCH DATA ALLIANCE  
**EUROPE**

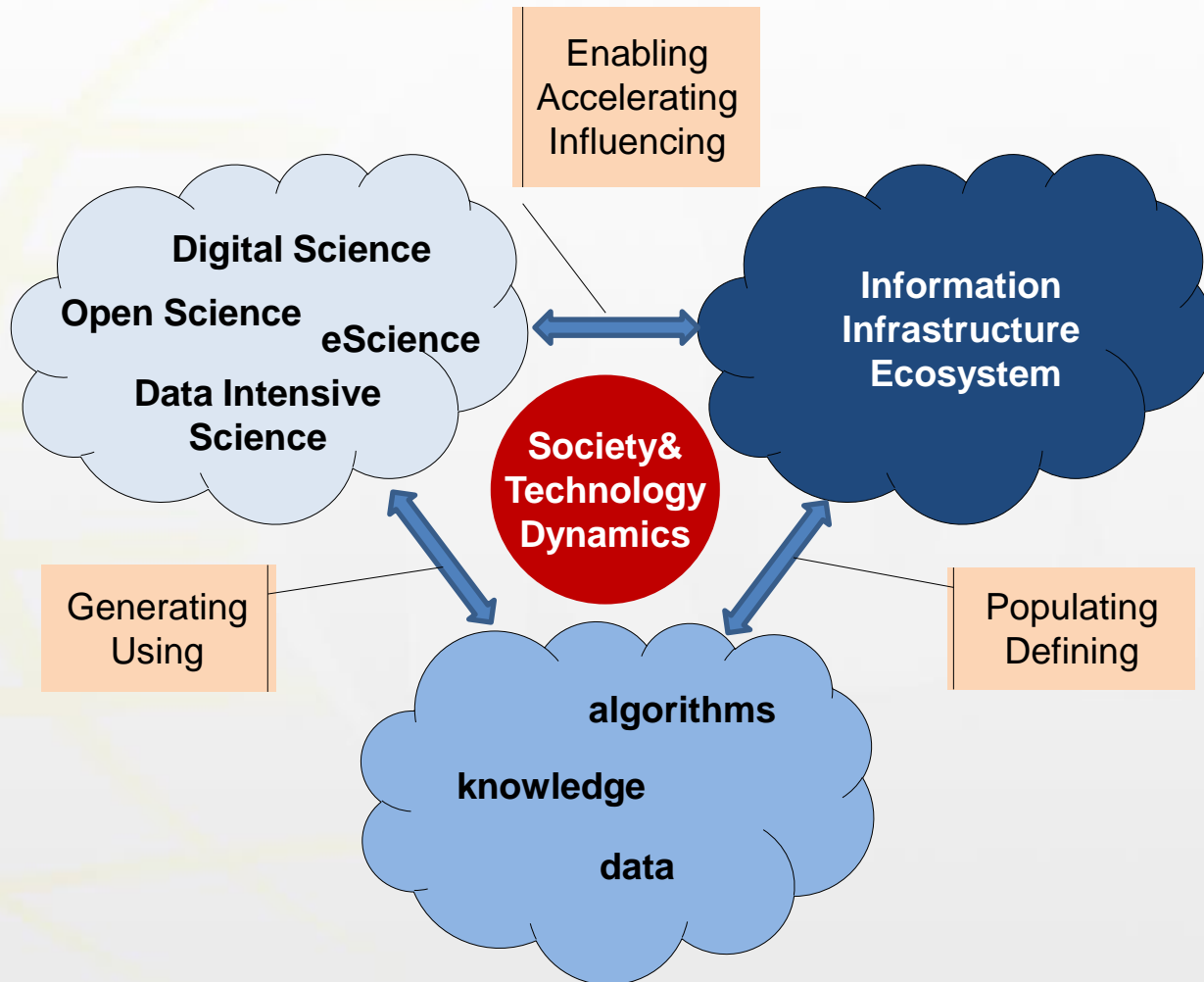
# Weaving the Data Fabric

An Ex-Practitioner View

Peter Wittenburg

Max Planck Computing and Data Facility

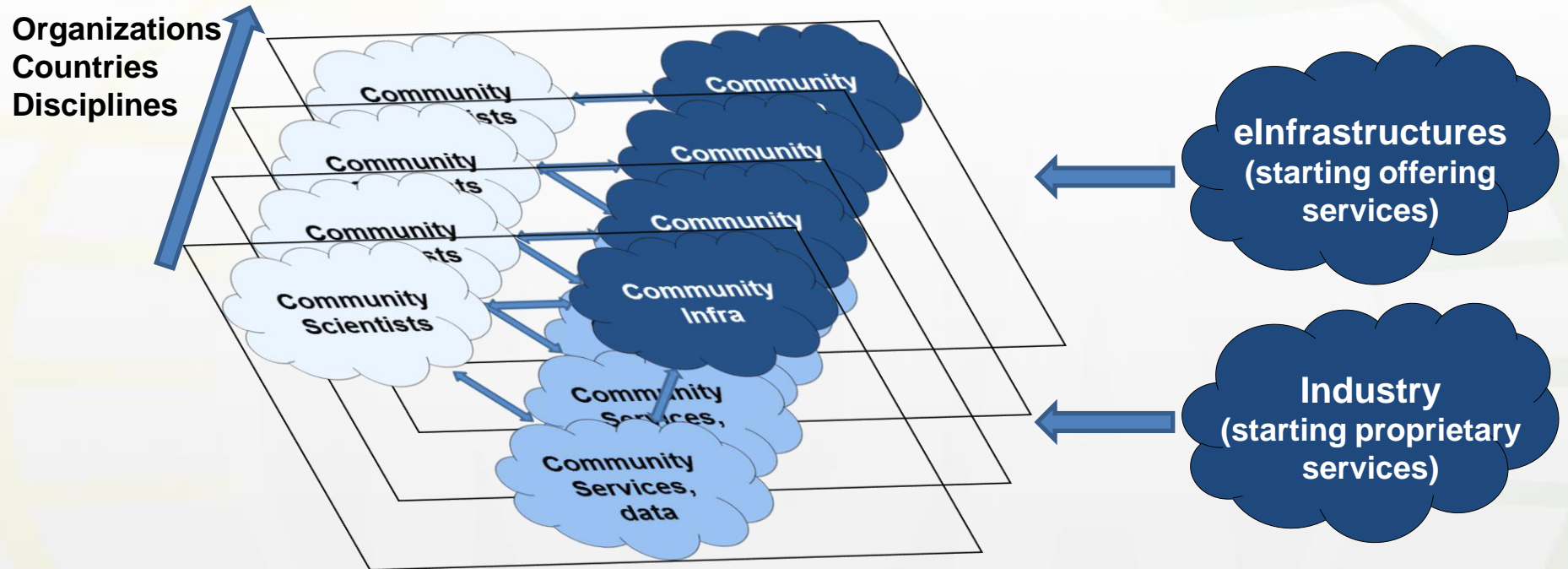
# Digital Change



## Key Terms

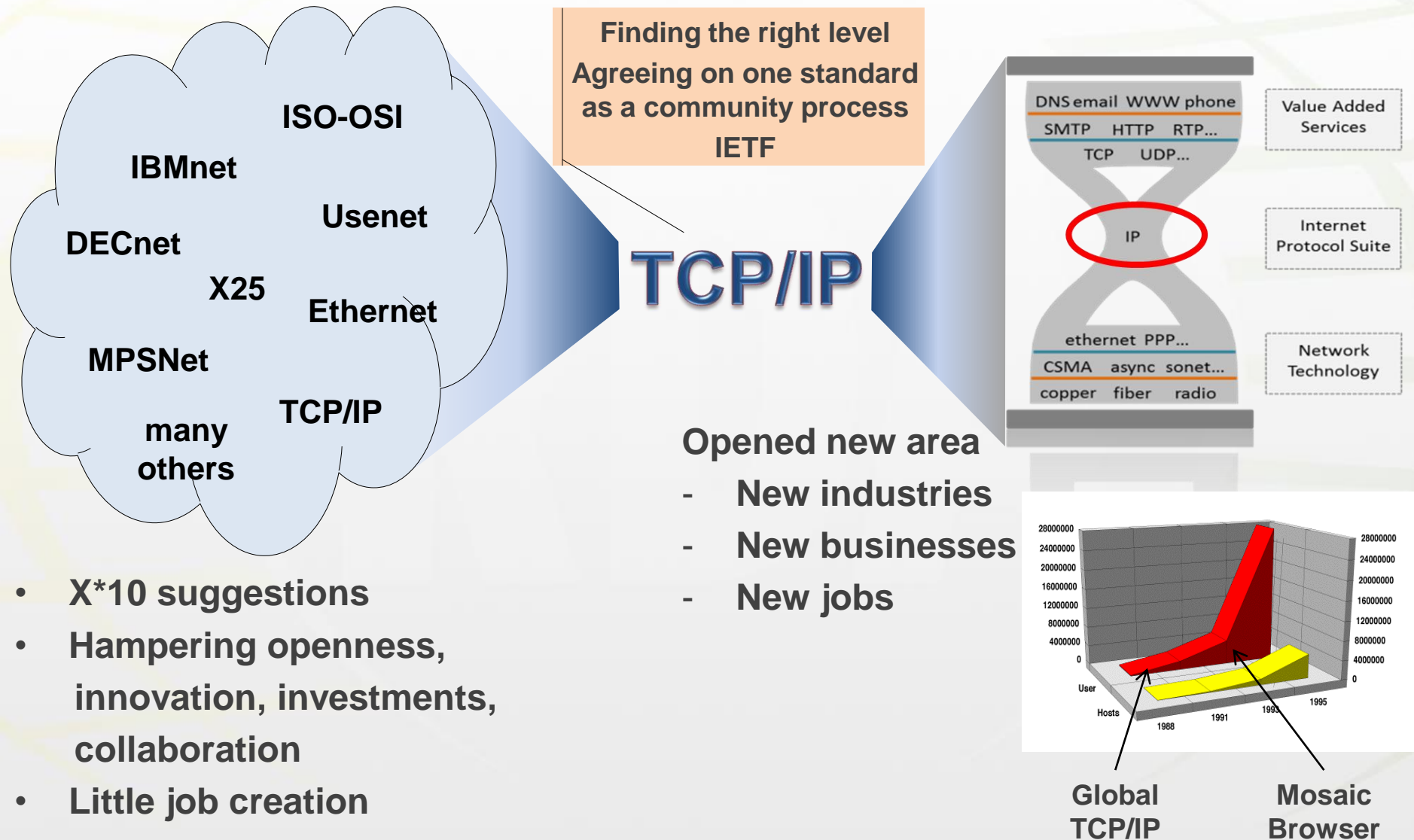
- Acceleration
- Innovation
- Competition
- Collaboration
- Openness
- Jobs

# State of Infrastructures



- ESFRI: much awareness raising in Europe, lots of young people trained, much testing of variety of approaches, identifying gaps in service landscape, etc.
- eInfra: starting to change towards service orientation, need more stable services, need clarification of costs
- All together: too many boundaries, too inefficient, too expensive, hampering software investment, hampering new businesses
- **And now? - lack harmonizations/reductions to make things efficient and sustainable**

# Internet Wayout

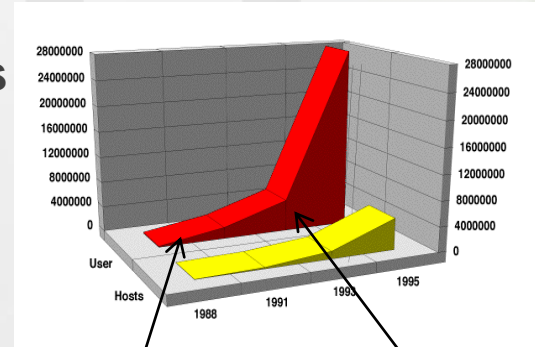


Finding the right level  
Agreeing on one standard  
as a community process  
IETF

**TCP/IP**

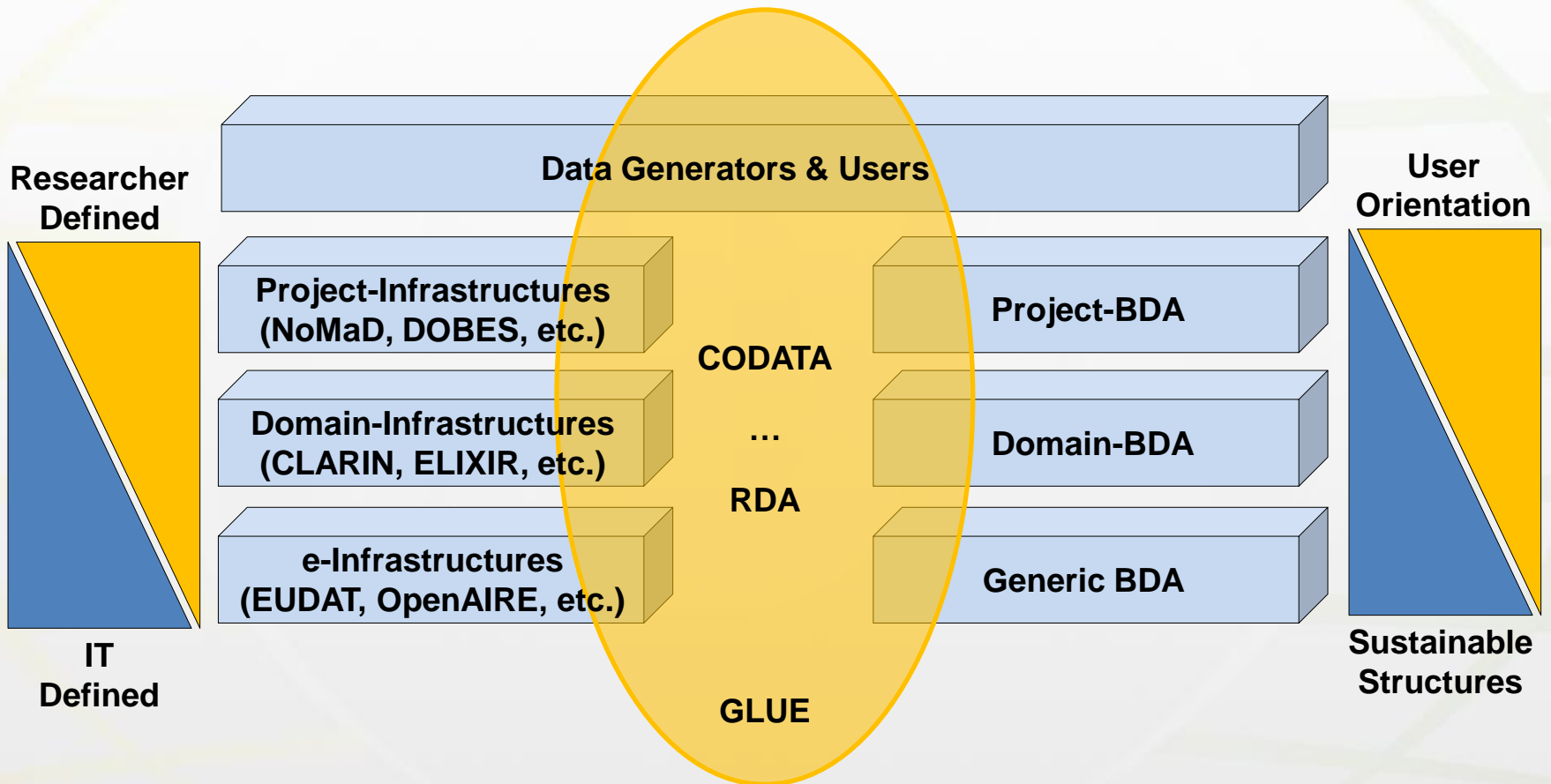
Opened new area

- New industries
- New businesses
- New jobs

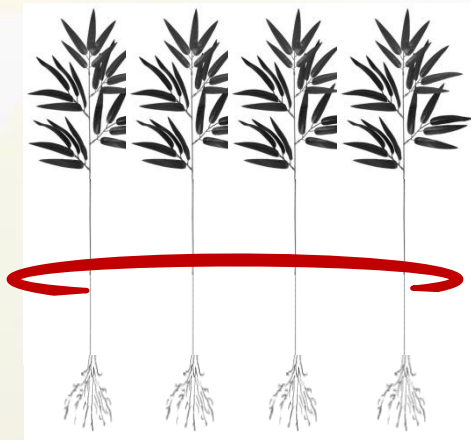


- X\*10 suggestions
- Hampering openness, innovation, investments, collaboration
- Little job creation

# Another view on Complexity



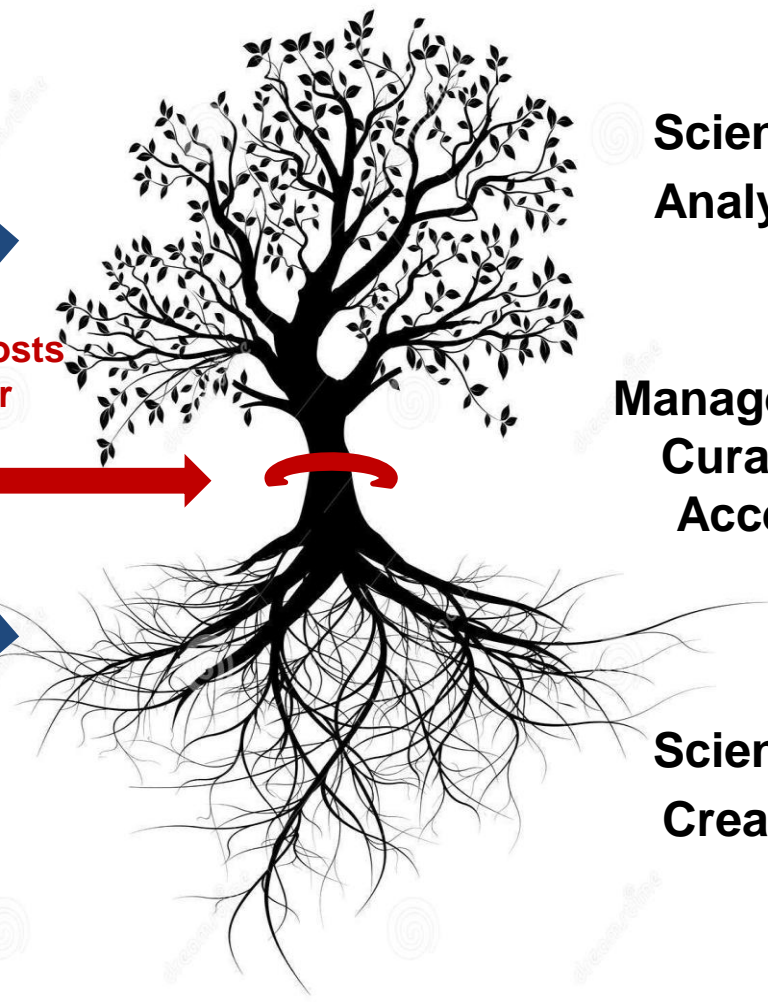
# Analysis of Infrastructures



Leave flexibility  
Even more opportunities

Reduce heterogeneity & costs  
Make solutions stronger  
Achieve sustainability

Leave flexibility  
Even more opportunities



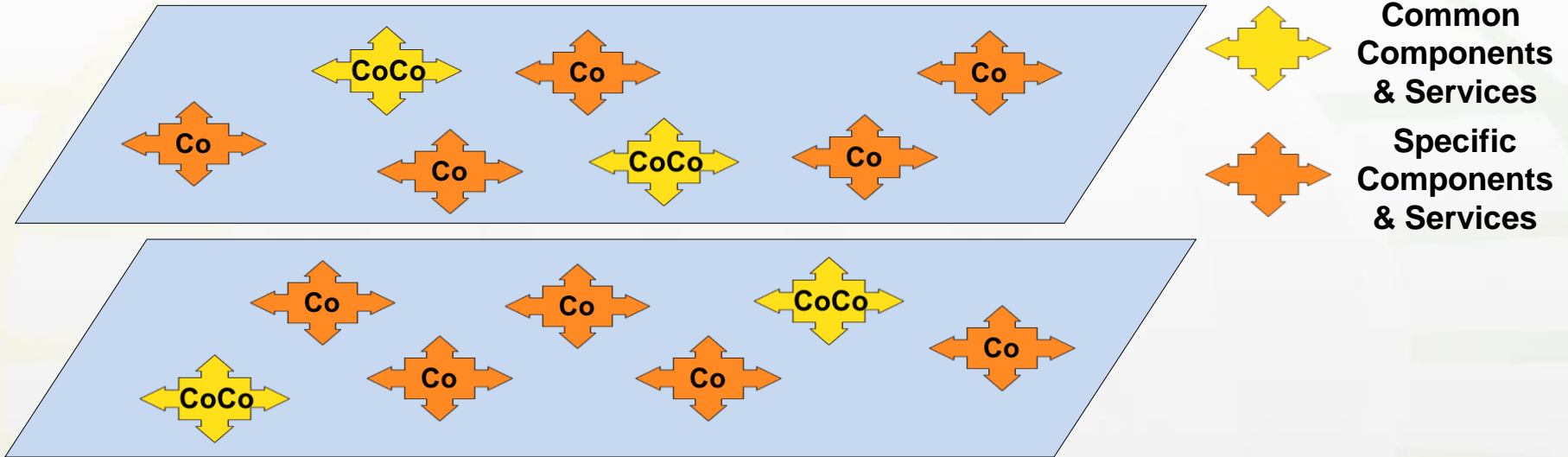
Scientific  
Analytics

Management  
Curation  
Access

Scientific  
Creation

PID, AAI, MD, WF,  
Registries,  
Repositories,  
meta-semantics,  
etc.

# Components & Data Fabrics

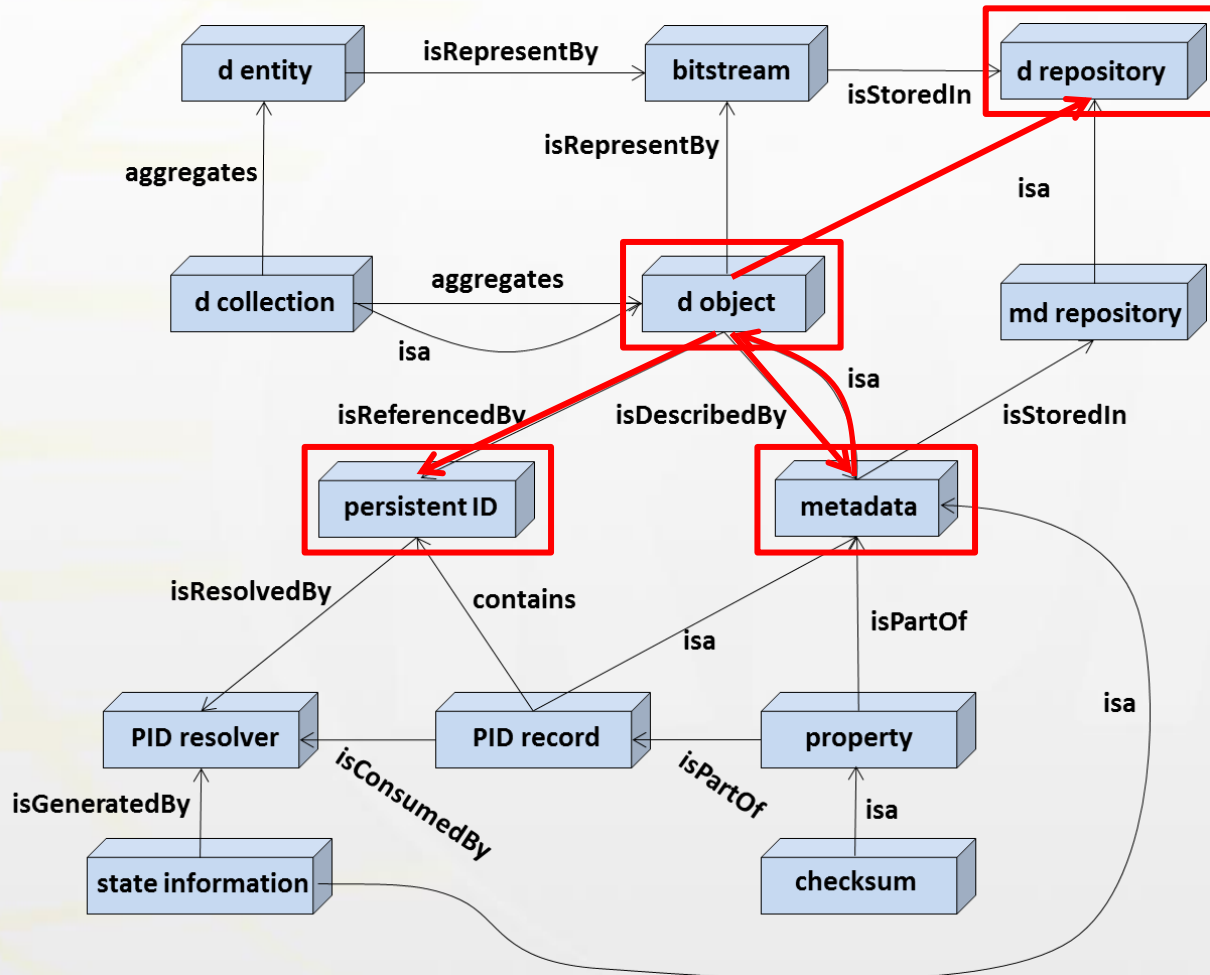


## Huge Task to solve:

- Identify and specify Common Components
- Recommend CoCo
- Put CoCo in place overcoming some resistance

**Who can steer this? - only cross-border & bottom-up organisation**

# Example: Data Foundation & Terminology



DFT core model is so simple.

It refers to a few obvious components.

Messages are on purpose very simple.

Need to learn to speak with one voice.

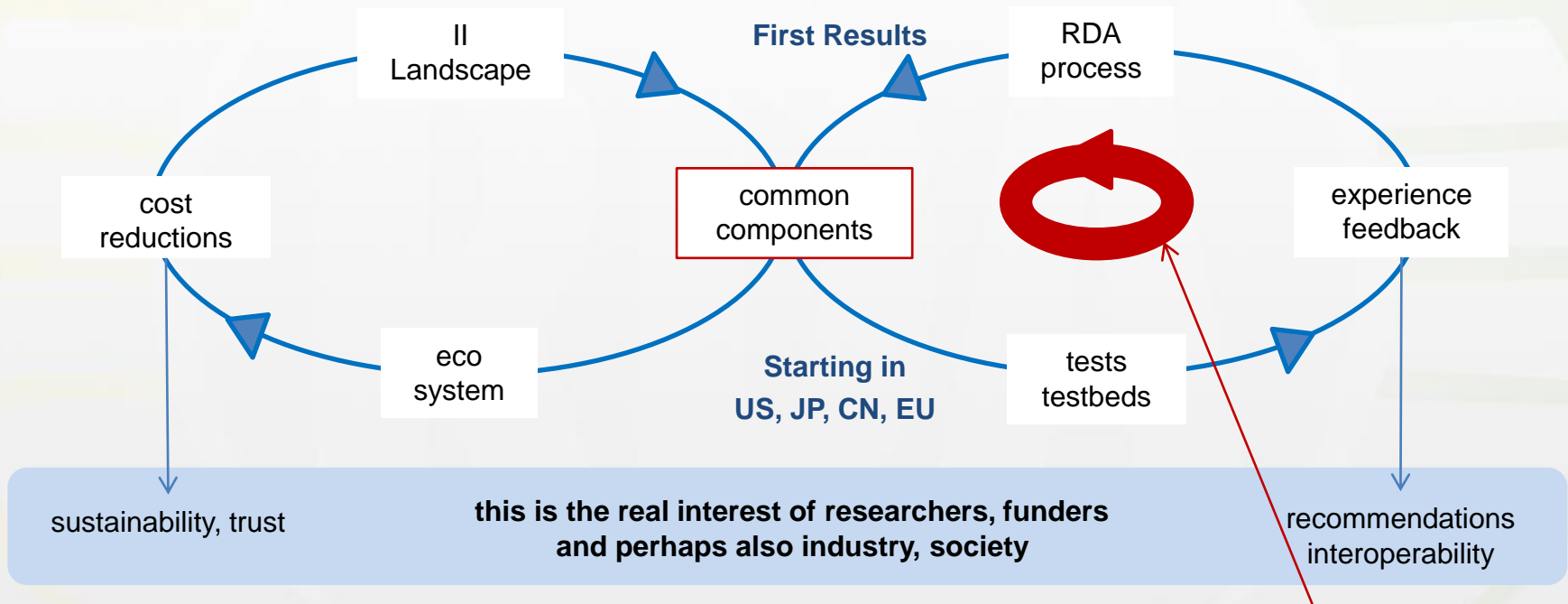
If all SW builders would adhere to it, we would have gained a lot!

Not to forget: PIT, DTR, PP, DCIT and other RDA results



# State of Discussions

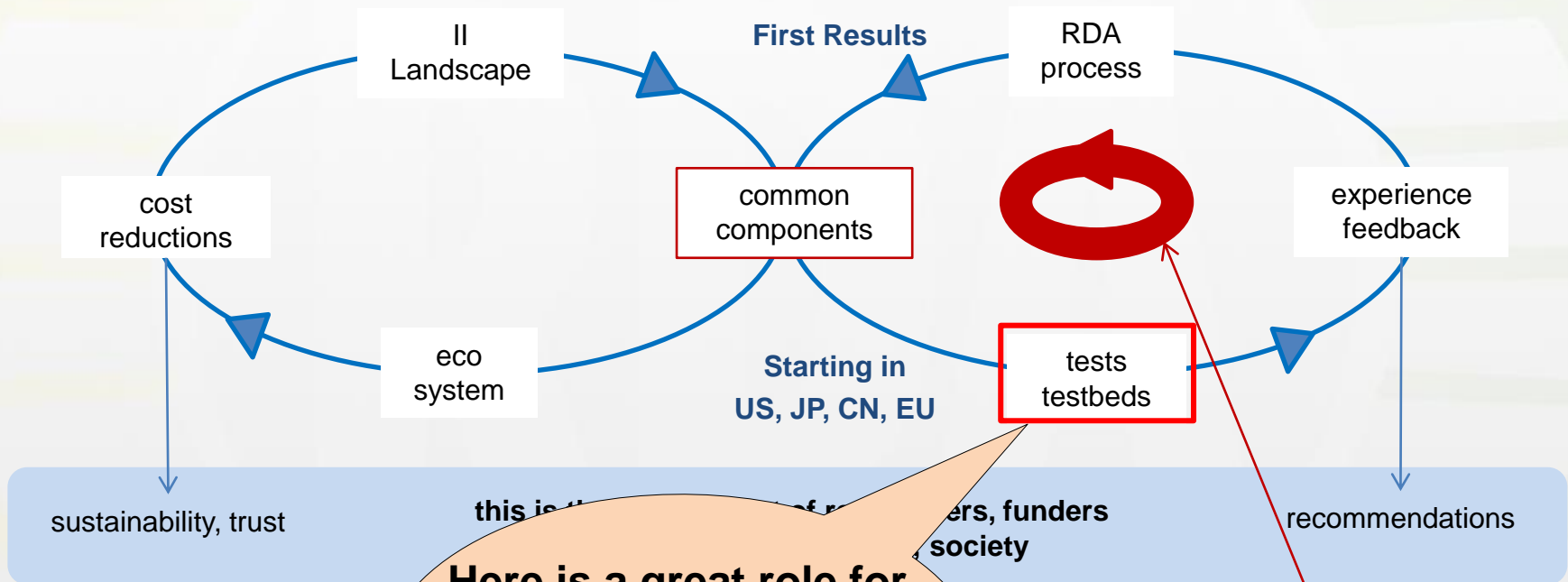
How to organize and speed up the process from Ideas to Recommendations?



- How much time do we have?
- How to communicate with funders?
- How to come to decisions?
- How to accelerate this cycle?
- How to increase density of debate?
- Papers from US, EU available

# State of Discussions

How to organize and speed up the process from Ideas to Recommendations?



- How much time do we need to accelerate this cycle?
- How communicate with participants to increase density of debate?
- How to come to decisions? Papers from US, EU available

# Need a Process NOW

- Make wide agreements recommendations
  - Can start with G8/FAIR/etc. principles
  - Another example: data objects to be assigned PIDs, associated with MD and bit sequences to be stored in certified repositories.
- Extract common components and specify their services from all tests and testbeds
  - Aggregate and analyse test results – all in open forum
- Can only be done cross-boundary (countries, disciplines, projects, etc.)
  - Need to overcome some “odd” behaviors
- Group of experts around RDA started collecting commons and provoking discussions -> Paris Document
  - Authors claiming no ownership on ideas etc. – socially very important
  - Received more than 30 contributions and use cases and more to come
  - Will discuss this in the Data Fabric IG sessions
  - Get engaged

# Value of RDA

- If there would be no RDA? – create one, cross-border (disciplines, countries, projects)
- If RDA is not doing well? – change it, but do it soon
- If RDA outputs are not ok? – increase density of testing, testbeds, adoption, aggregate experience and feedback to RDA process
- If RDA is not fast enough? – accelerate this cycle by even more testing
- RDA EU is pushing a lot with its **collaboration** call and its **Senior/Junior support team** – make use of it
- RDA EU, EUDAT, EGI, OpenAIRE, LEARN, EDISON to bundle forces in training and giving advice
- Can RDA make recommendations? – yes, tested results show the way
- Can RDA take decisions? – probably not, it must stay a neutral place to discuss alternatives, technologies etc.

# Thanks for your attention.



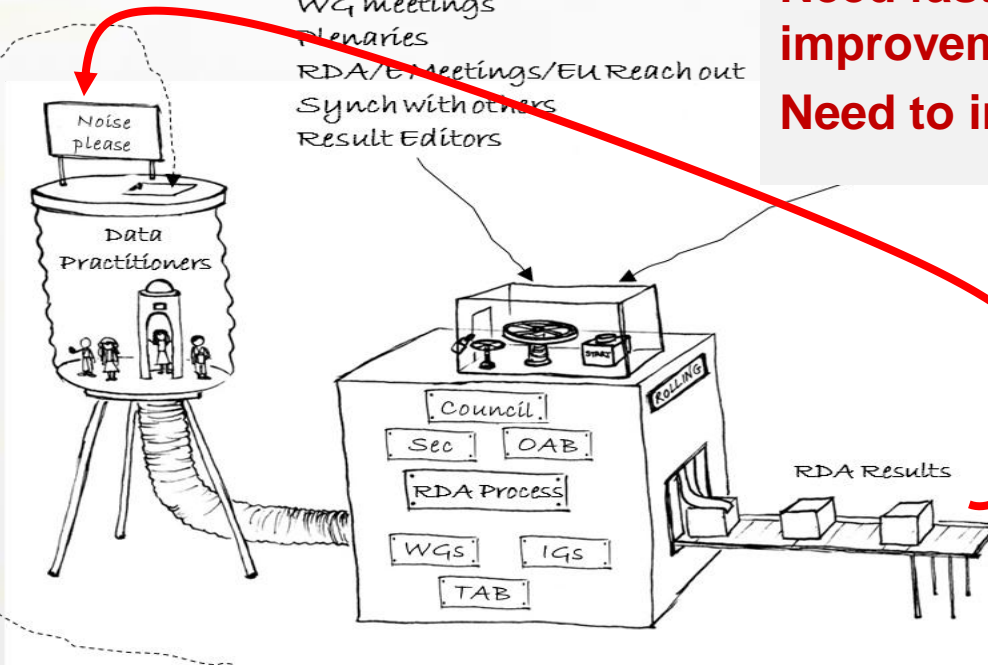
John Wood: Live in an exciting time, it can all go horribly wrong.

# RDA Machine

**Need fast feedback and improvements.**

**Need to increase discussion density.**

Support process  
Council, OAB, TAB, Sec, Chairs  
WG meetings  
Plenaries  
RDA/E Meetings/EU Reach out  
Synch with others  
Result Editors



Result Dissemination  
Storing/Publishing Results  
Offering Results  
Gather Impacts  
Training about results  
Evangelization and Interaction  
Help to implement results  
Help to adapt to results  
Help to adapt software

The RDA-RDA/E Machine

Phase II: How to push the bottom-up work and uptake