New User Communities

Michael Grønager, PhD

GN3 Innovation Workshop
Copenhagen, October 11th 2011

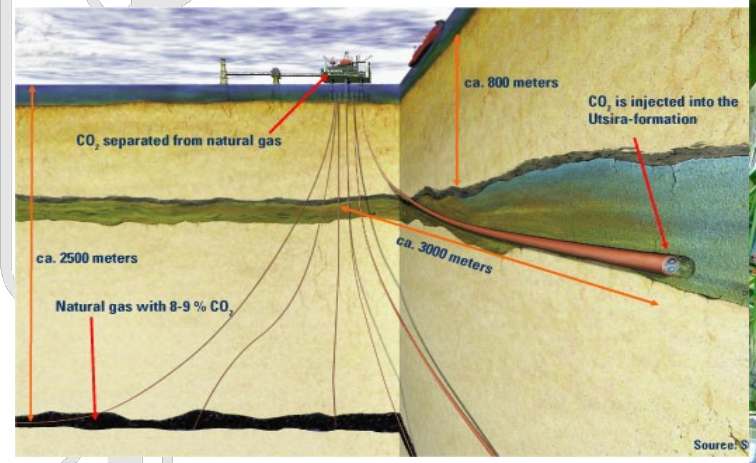
NDGF

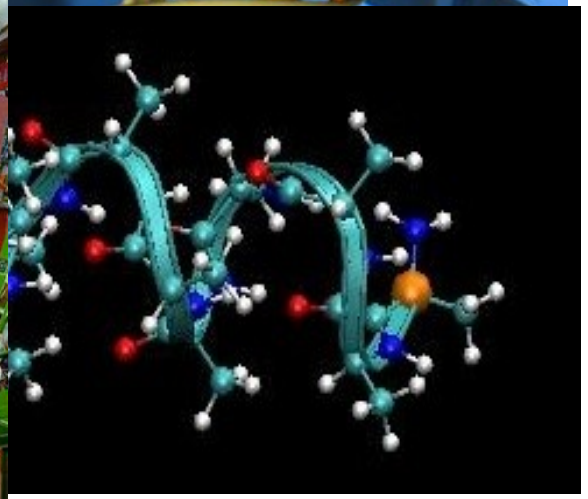
Background

- Considerations from NDGF 2006-2011
- NDGF @ NORDUnet A/S
  - Operating the Nordic CERN Tier-1
  - Supporting the Nordic CERN Tier-2s
  - And New User Communities
- A common infrastructure for e-Science ?

NDGF - Dist. Computing Collab.

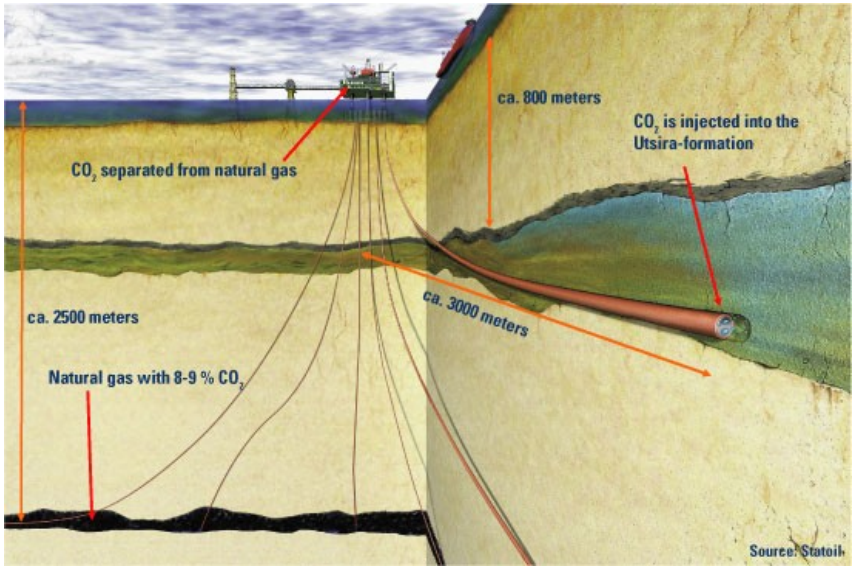
- eScience for Nordic Virtual Research Communities
  - Nordic WLCG Tier-1
  - BioGrid community grid
  - CO_2 sequestration community grid
  - Computational Chemistry grid
  - Material Sciences





NDGF Case I: CO2 Seq.

- Achievements
  - Make it easy for the user to get on the grid
  - Hiding the grid in an application server
  - Grid support for code compilation
- Some success – a lot of huge runs in 2008
  - Mainly access to compute power
  - Hardly any need to share data
  - Was grid really the right way ?



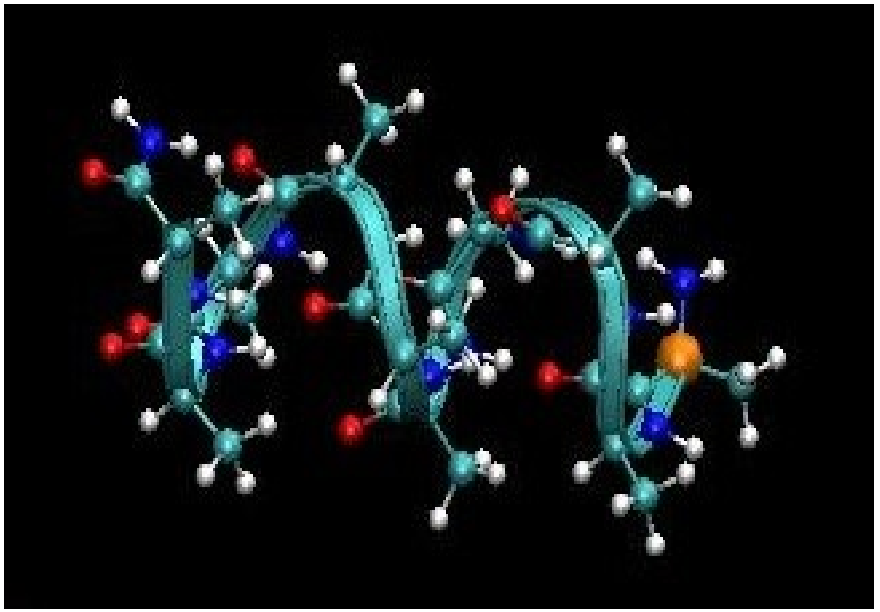
NDGF
Nordic DataGrid Facility

NDGF Case II: Bio-Info.

- Achievements:
  - Dynamic databases
    - Auto updating of these
  - Out of the box MPI enabled bio applications
    - BLAST, HMMER, "R", etc...
- In active use today
  - However, massive use is lagging
  - Integrated into the storage system
  - Not coupled with European efforts
  - Community not fully consolidated?

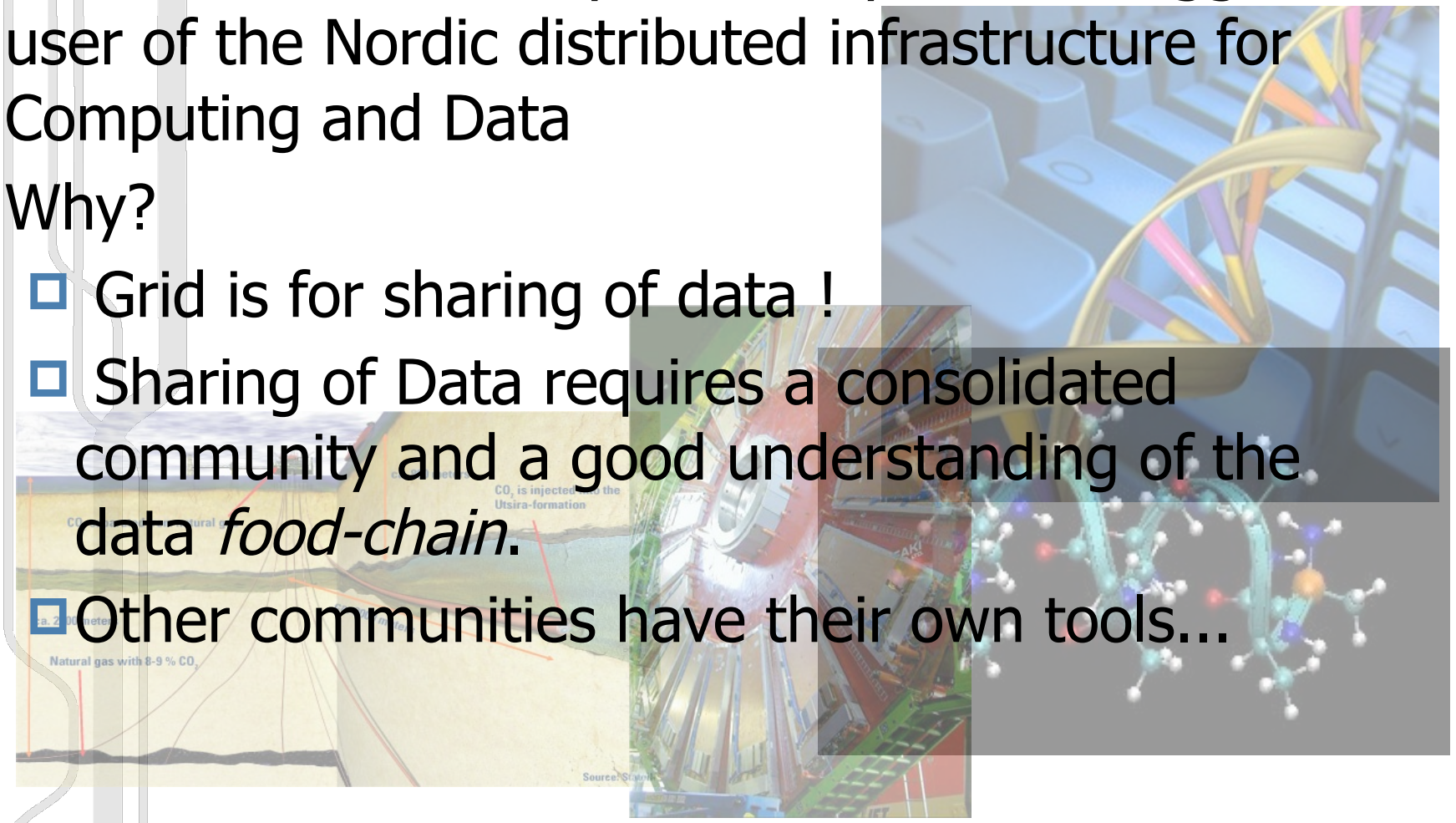
NDGF Case III: Chemistry

- Achievements:
  - Auto install based on code snapshots
    - Development based on the CO2-seq. Project
  - Further improvements for "MPI on Grid"
- Hardly in use today
  - No need to share data
  - Is grid right way here ?



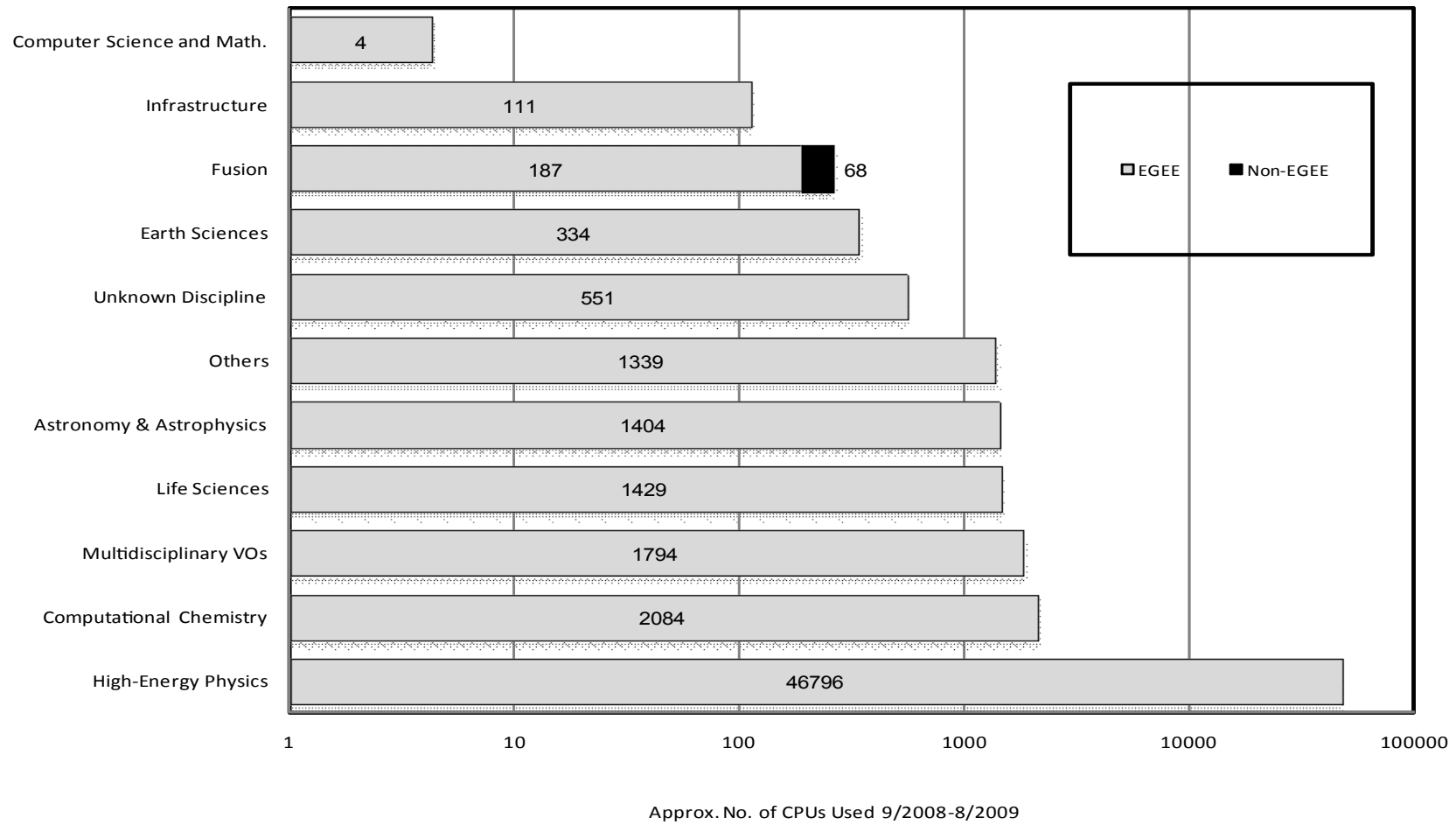
NDGF - Dist. Computing Collab.

- The WLCG community is still, by far, the biggest user of the Nordic distributed infrastructure for Computing and Data
- Why?
  - Grid is for sharing of data !
  - Sharing of Data requires a consolidated community and a good understanding of the data *food-chain*.
  - Other communities have their own tools...

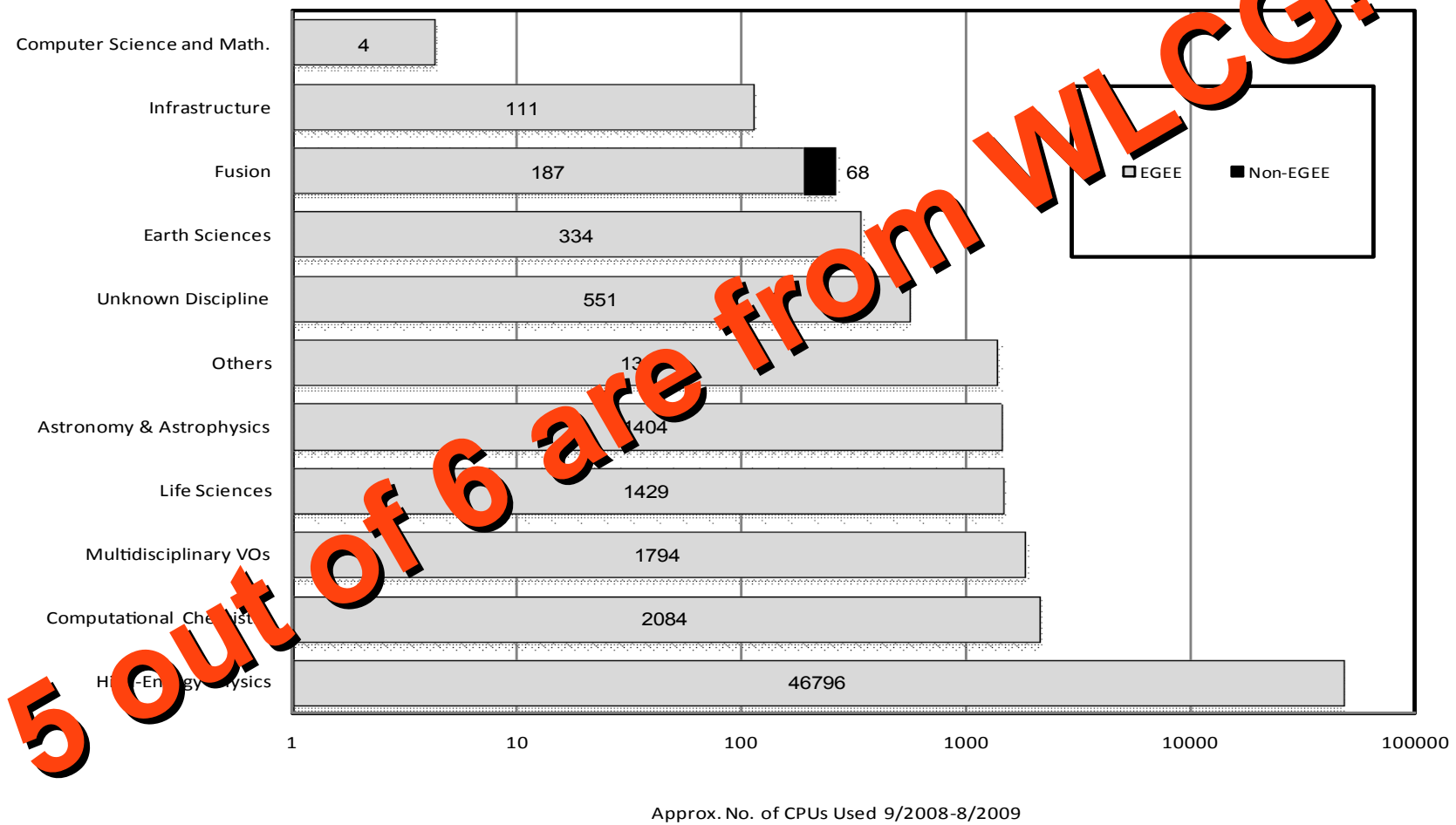


NDGF
NORDIC DATAGRID FACILITY

The EGI Users of 2010



The EGI Users of today



5 out of 6 are from WLCG!

Virtual Research Communities?

- ESFRI:
  - ELIXIR
  - ESSS
  - ICOS
  - LIFEWATCH
  - CLARIN
  - BBMRI, EPOS, EISCAT-3D, …

A common InfraX for eScience

- Enabling X disciplinary research
- Rationalization

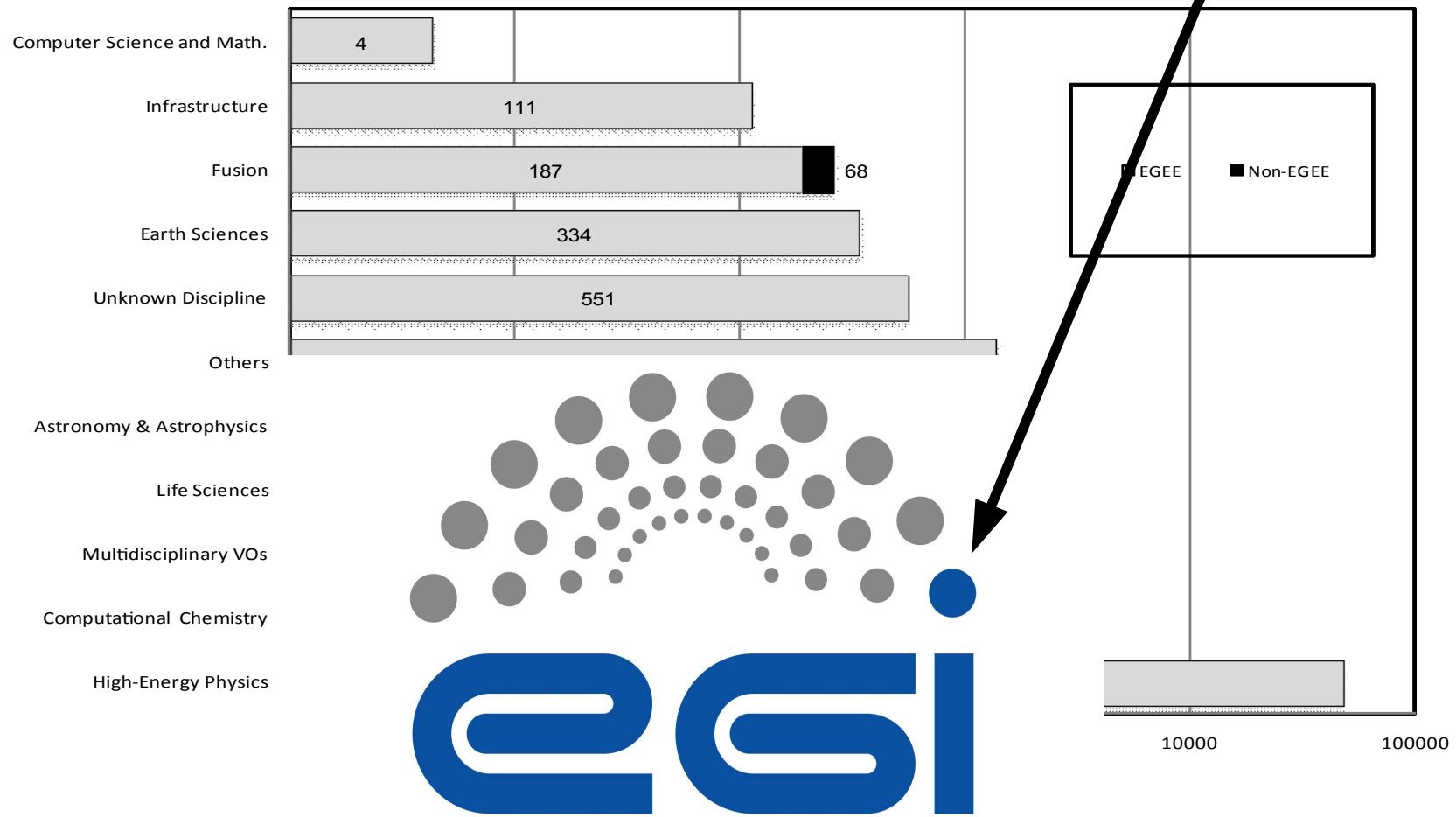
- However:
  - Each of the ESFRI projects have their own Distributed Computing Infrastructure
    - WLCG alone have several

- The idea of a common infrastructure for eScience seems unrealistic

A common InfraX for eScience

- New goals ?
  - Targeted Interoperation between the infrastructures.
  - A lower level infrastructure
    - Clouds? - might rationalize operation but will not enable x-disciplinary res.
  - Use of and contribution to existing open source projects
    - Long perspective ?

Thanks!

The EGI Users of today

Counting communities has started ;-)