Netherlands eScience Center

Introduction
&
Climate applications

Prof. dr. Wilco Hazeleger

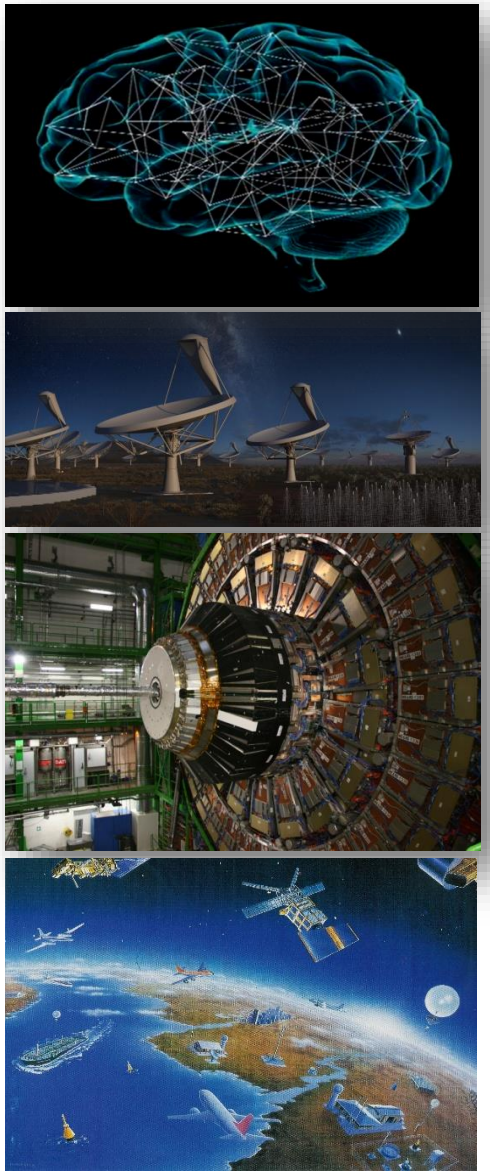
netherlands eScience center

by SURF & NWO

Todays big challenges

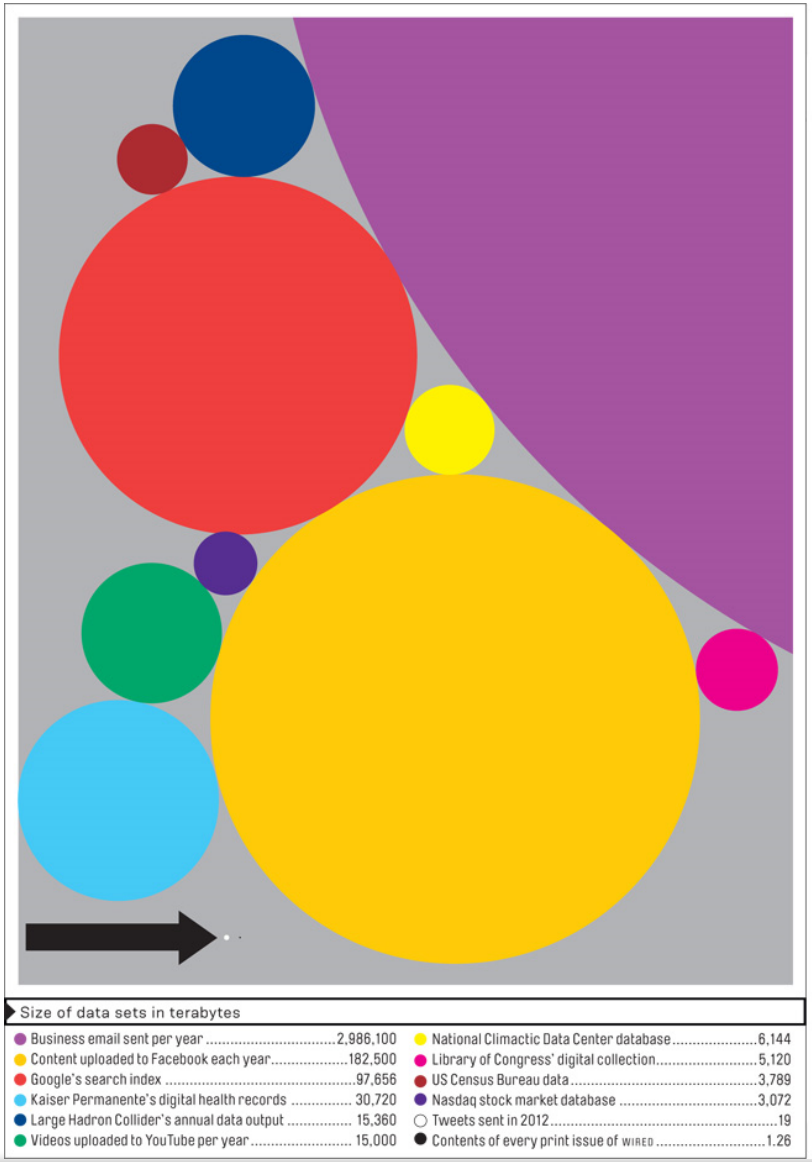
- **Big science often driven by big societal questions**
  - **increasingly multi, inter, transdisciplinary** ***(circular economy, healthy society, sustainability,..)***
  - **nearly all have an ICT or data component**

- **Science itself big and/or complex**
  - **Big data, compute, analytics**
  - **Big projects**
  - **Interactions between systems and scales**



Data-driven science

- **Miniaturization & parallelization**
- **Deployment of remote sensor networks**
- **Digitization of information & society**
- **Integration of instruments to the net**
- **Capacity to compute, store and share data**

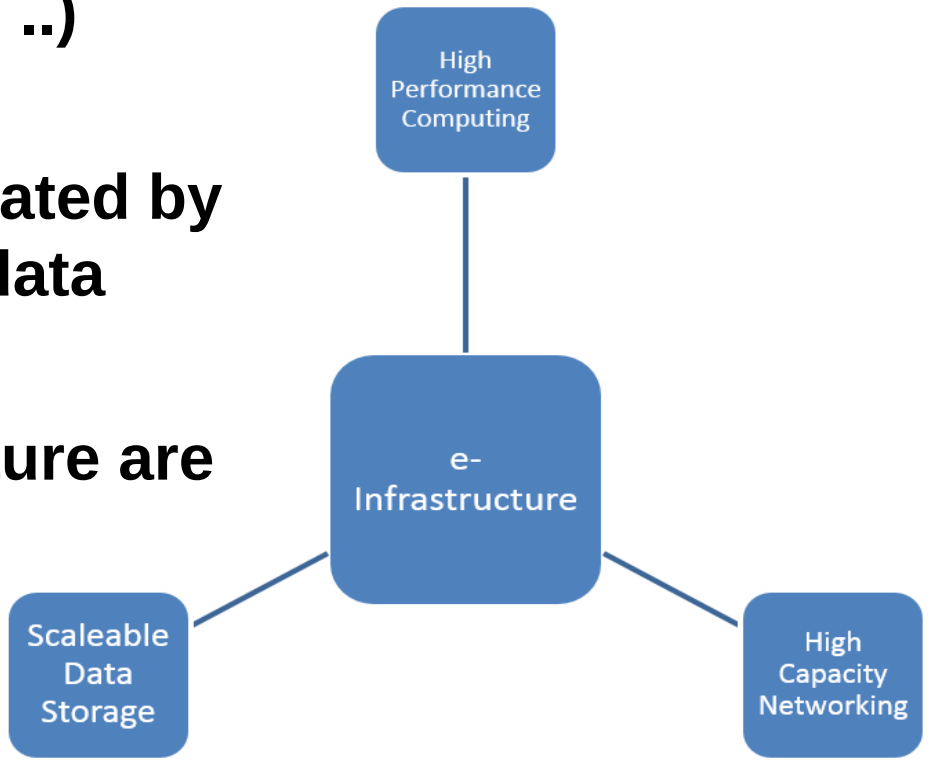


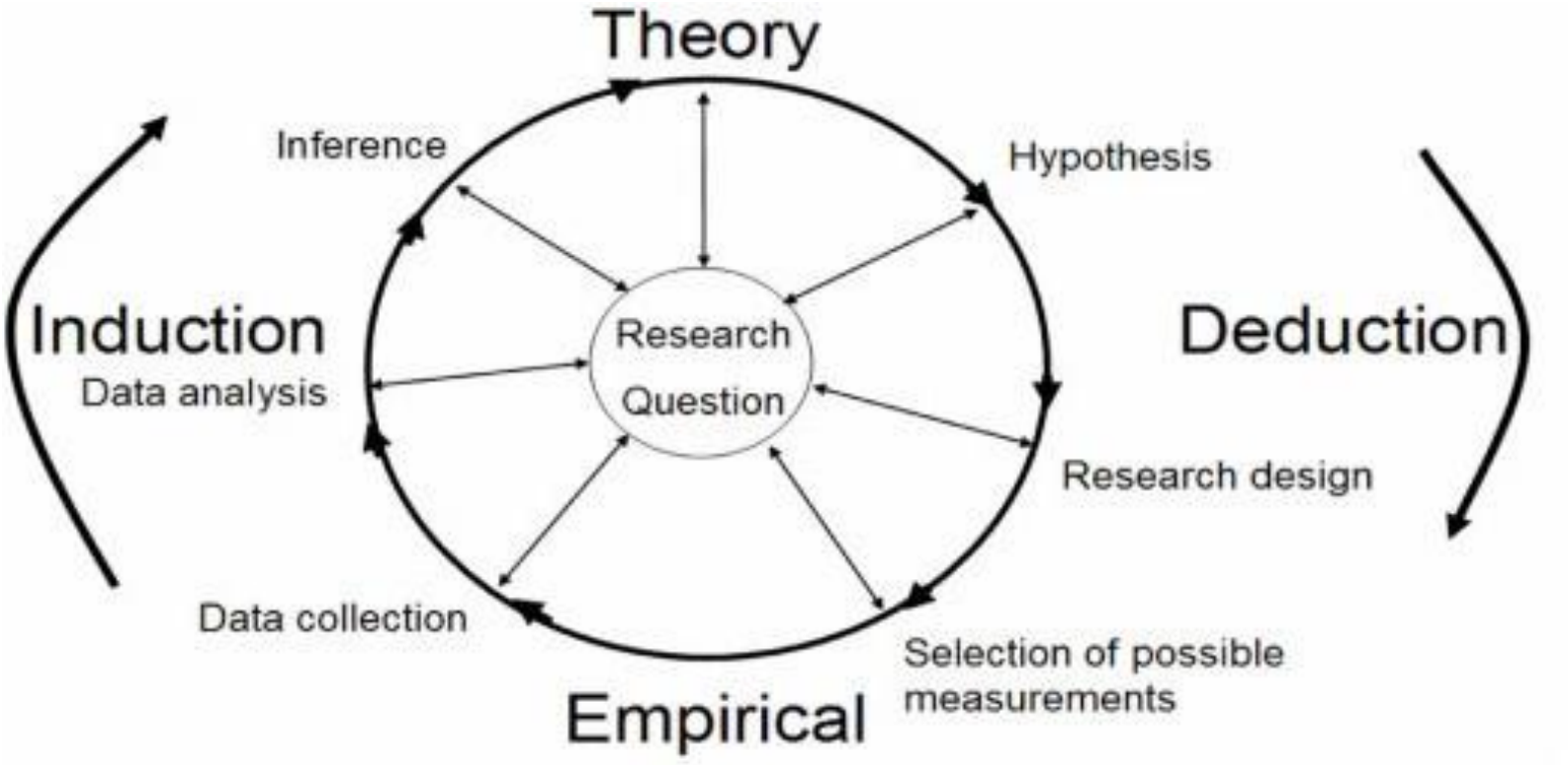
e-Infrastructure

- **Historically, scientific domains were defined by their unique tools (microscopes, telescopes, ..)**

- **Modern research is dominated by generic activities around data**

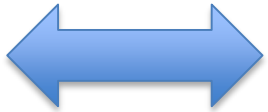
→ Computers & e-Infrastructure are the common component




Theory
Inference
Hypothesis
Induction
Data analysis
Research
Question
Deduction
Research design
Data collection
Selection of possible
measurements
Empirical

Netherlands eScience Center

advanced ICT, computer, data science



research

SURF

NWO

How we work - overarching

- eScience Research Engineers
- Fund projects (in cash and in kind)
- Public private collaborations (PPS)
- Coordinate (internationally: PLAN-E)
- Generic eScience Technology Platform
  - Overarching software
  - Knowledgebase & training

NLeSC: expertise center

Strategic Expertise Areas

Optimized Data-Handling

Big Data Analytics

Efficient Computing

Priority Domains

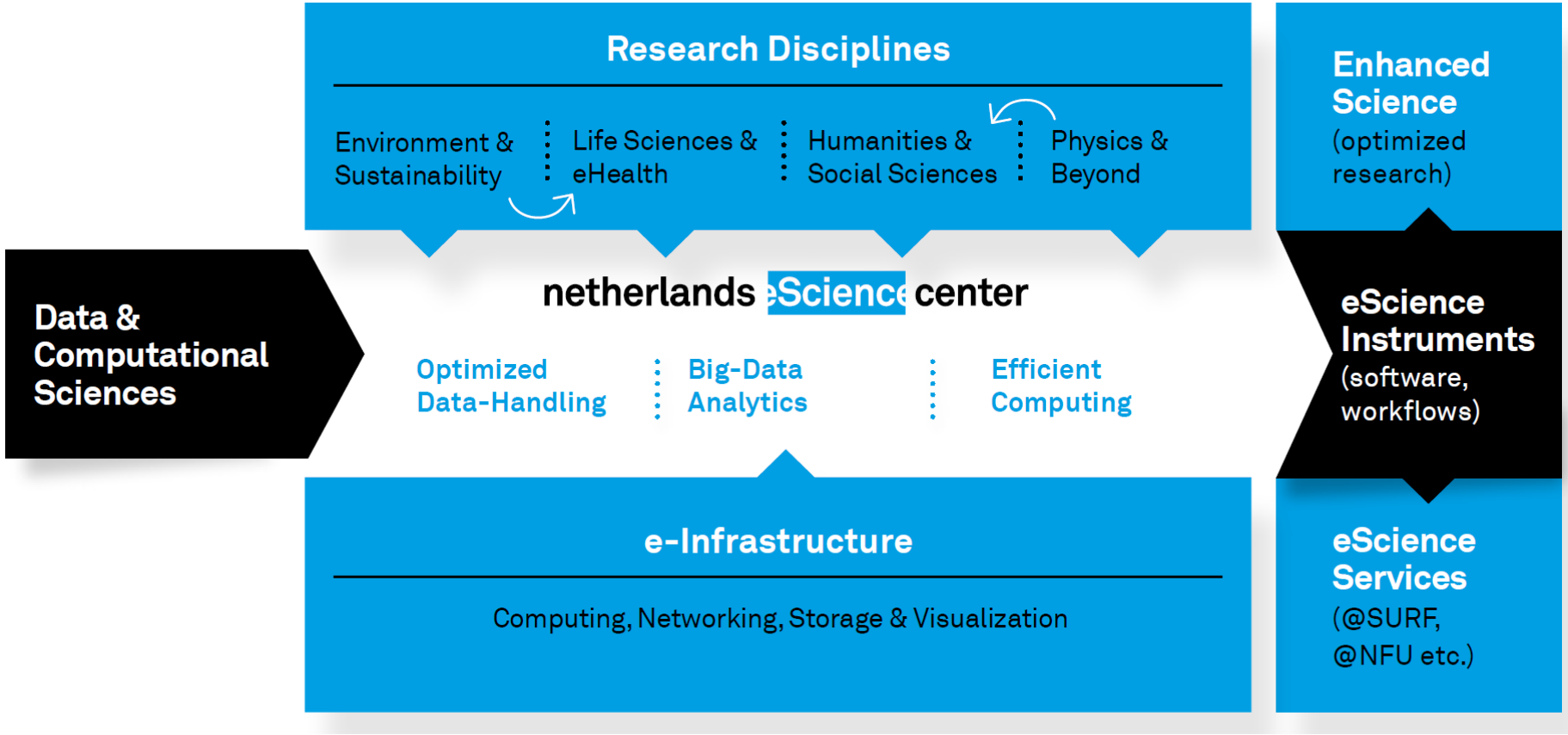
Life Sciences & eHealth

Humanities & Social Sciences

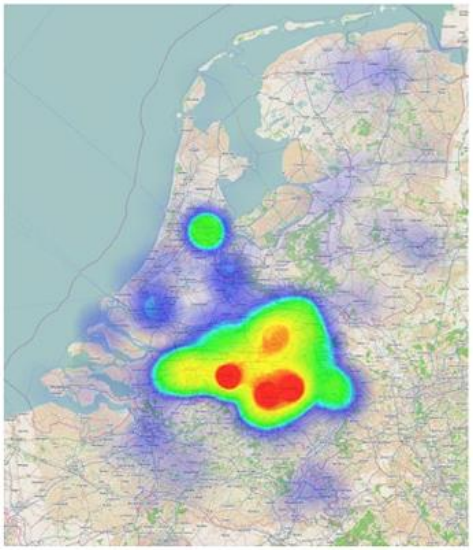
Sustainability & Environment

Physics & Beyond

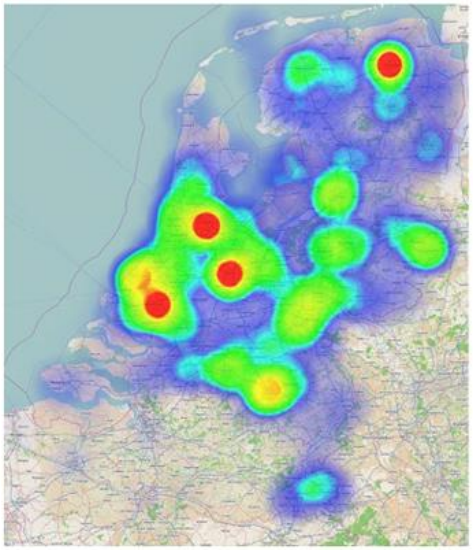
Bridging Research & e-Infrastructure



TwiNL



Dissemination of the Brabant dialect term "hedde" in January 2013.

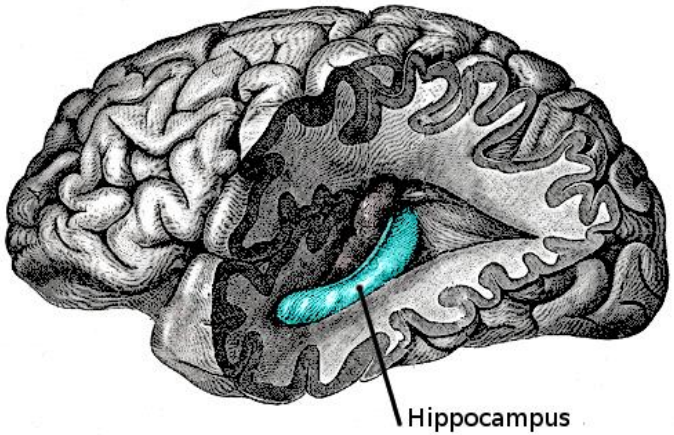


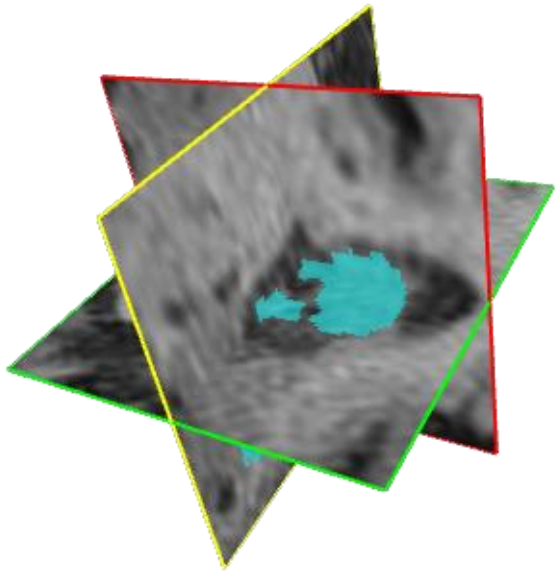
Dissemination of the word "alweer"
Time period: 22 January 2013 - 22 February 2013
(largely used in the Netherlands)

Dissemination of the word "weeral"
Time period: 22 January 2013 - 22 February 2013
(largely used in Flanders)

Dialect words out of ~5 million Dutch tweets a day

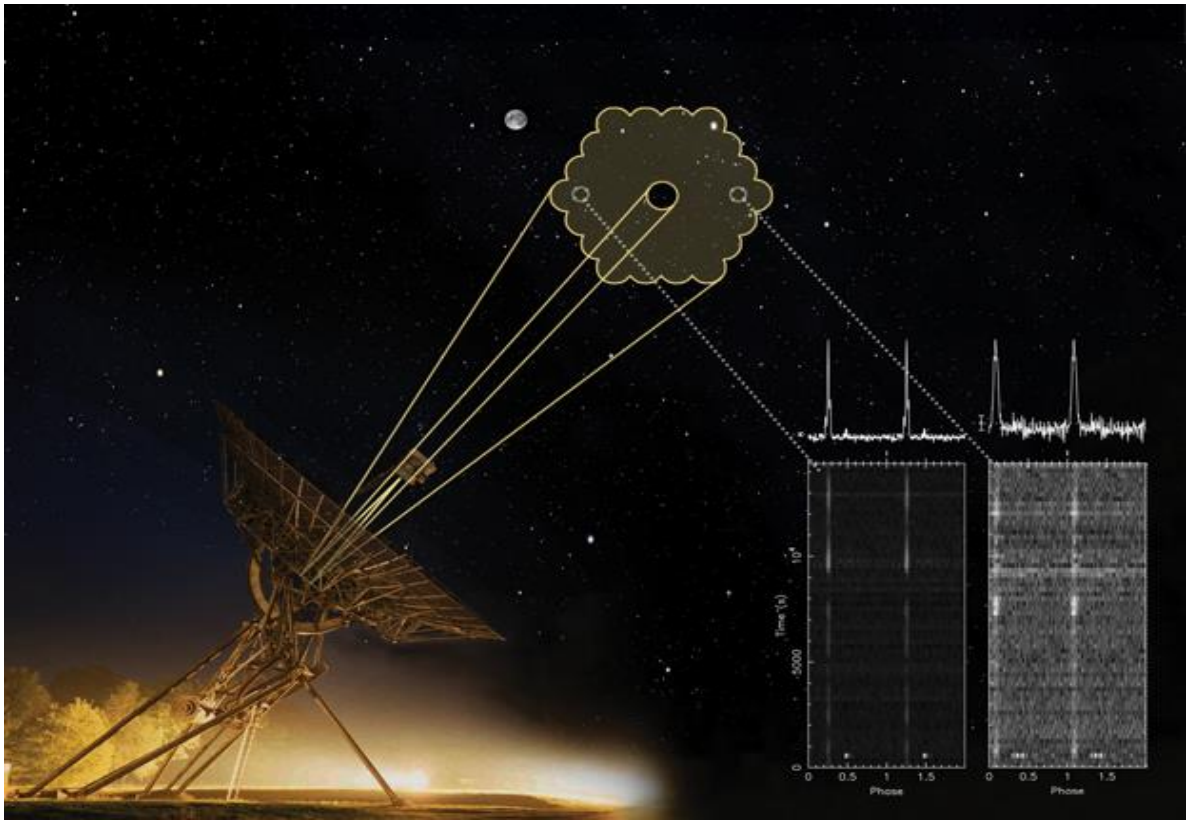
New markers for human health

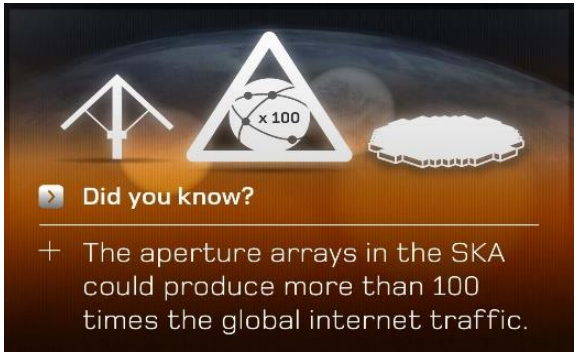




Mapping our world in 4D

Big data for the Big Bang





Embodied emotions

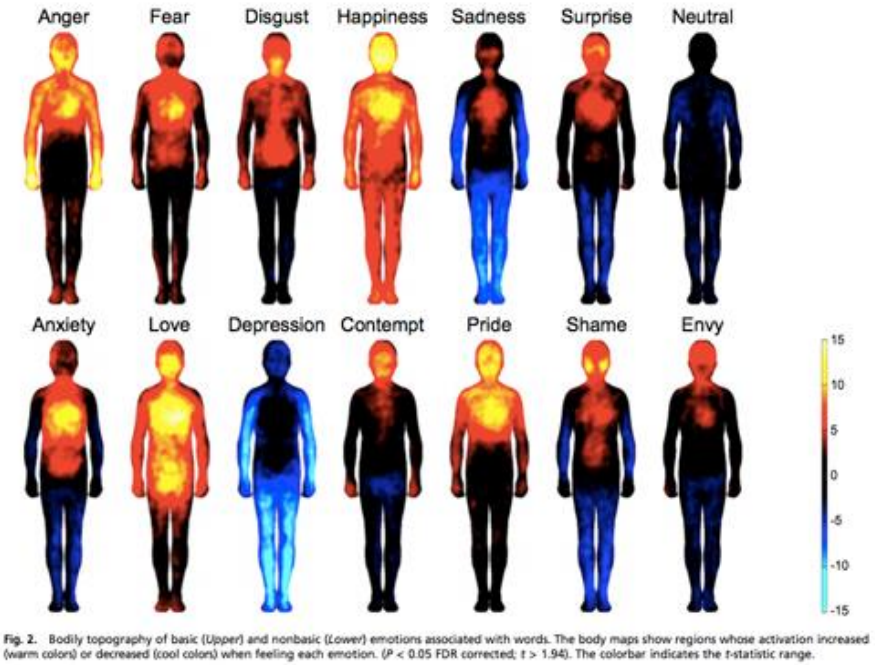
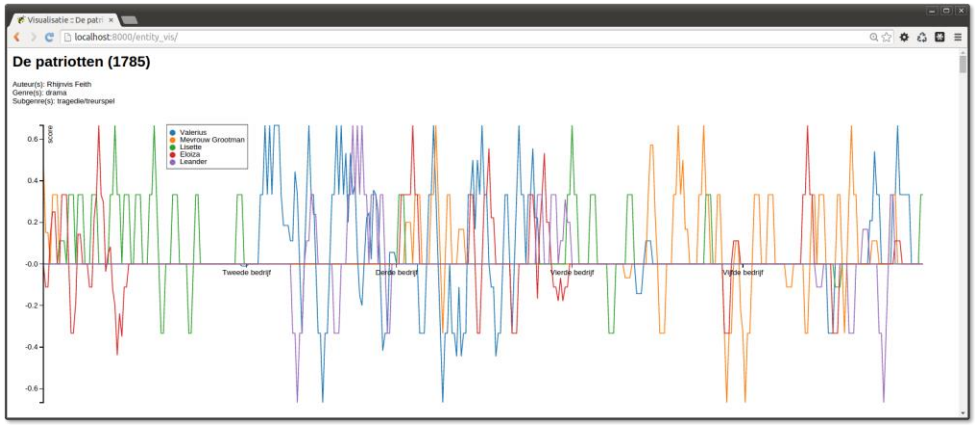
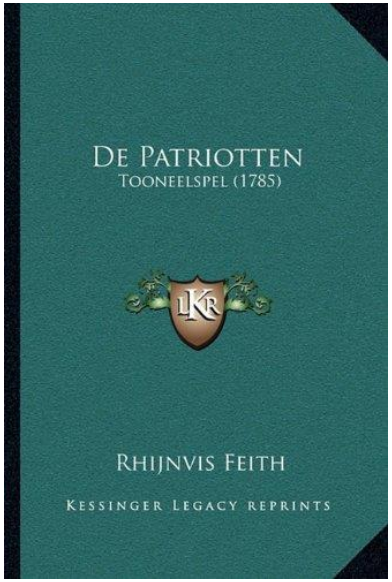
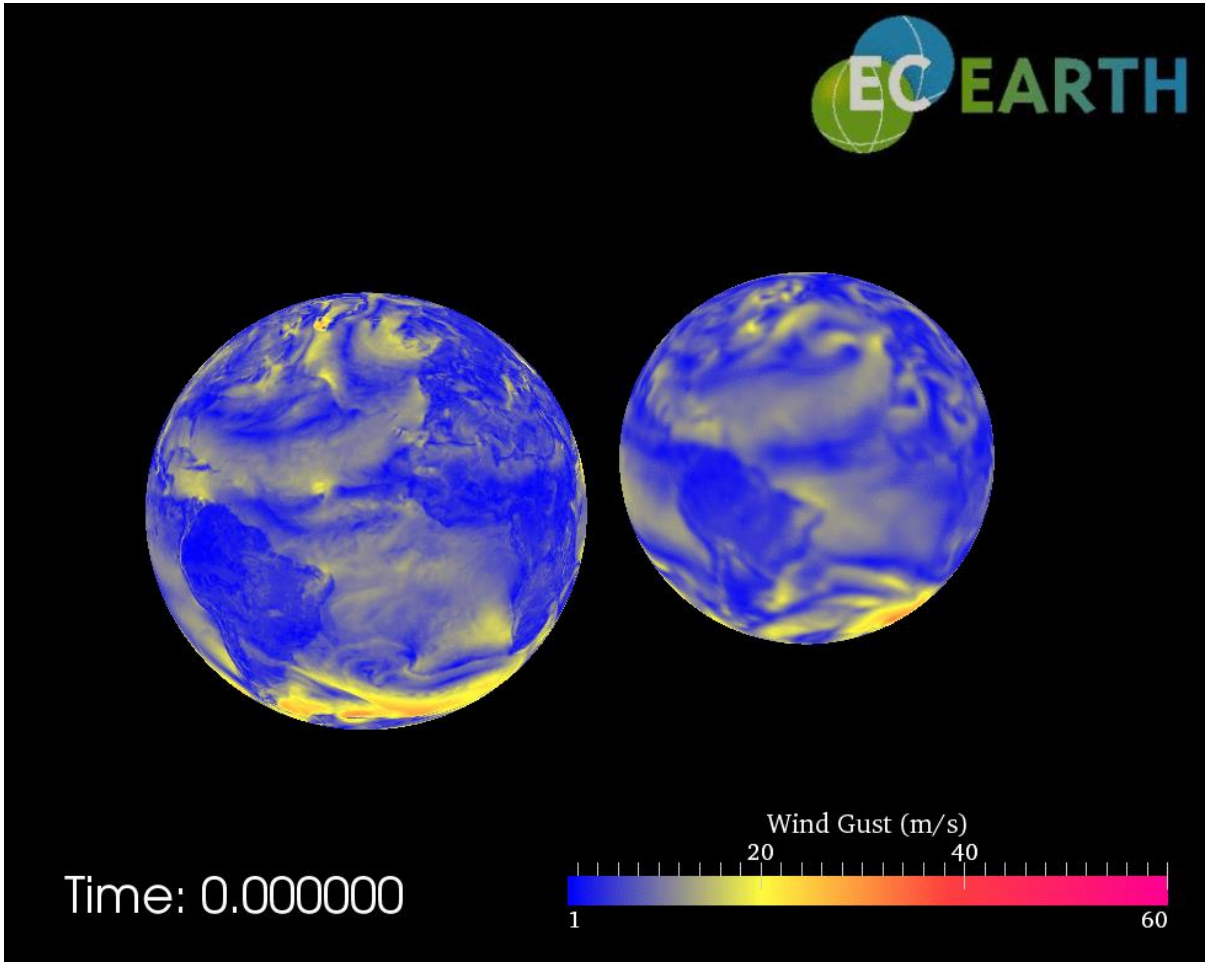


Fig. 2. Bodily topography of basic (Upper) and nonbasic (Lower) emotions associated with words. The body maps show regions whose activation increased (warm colors) or decreased (cool colors) when feeling each emotion. ($P < 0.05$ FDR corrected; $t > 1.94$). The colorbar indicates the t-statistic range.



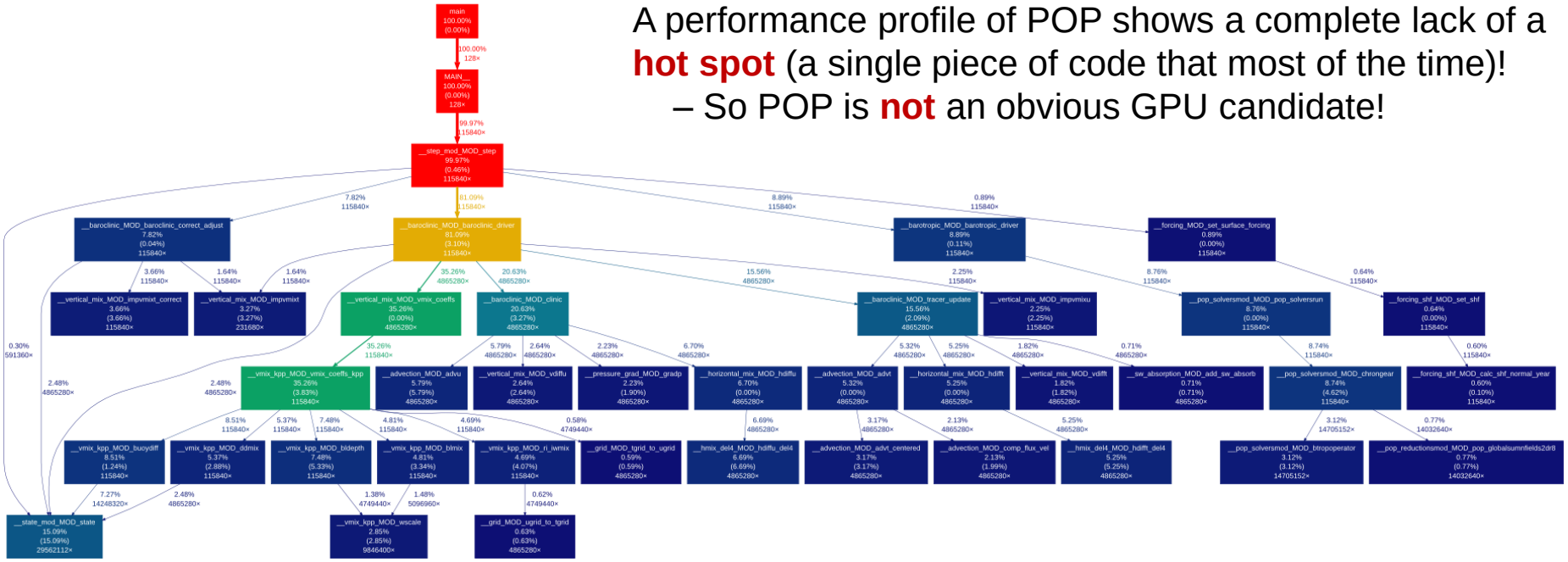


High resolution coupled climate modelling



e.g. H2020 PRIMAVERA & HIGHRESMIP

Are ocean models suitable for GPUs ?

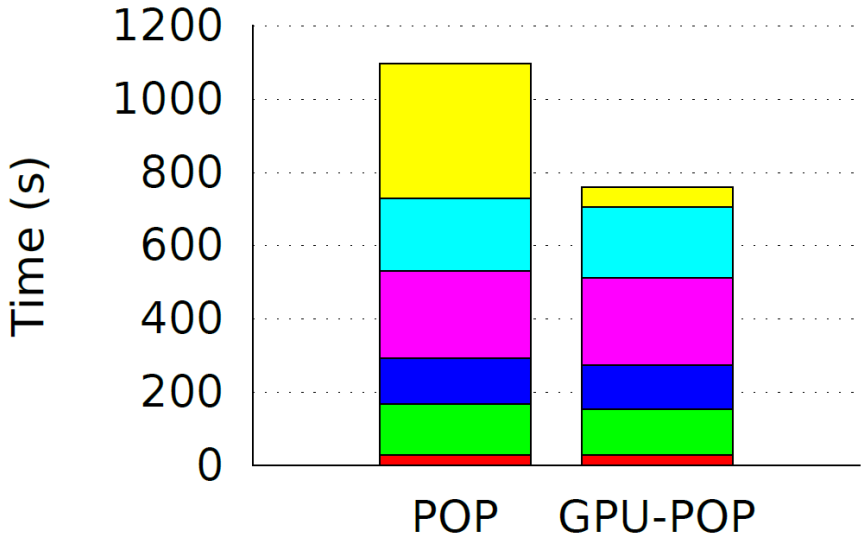


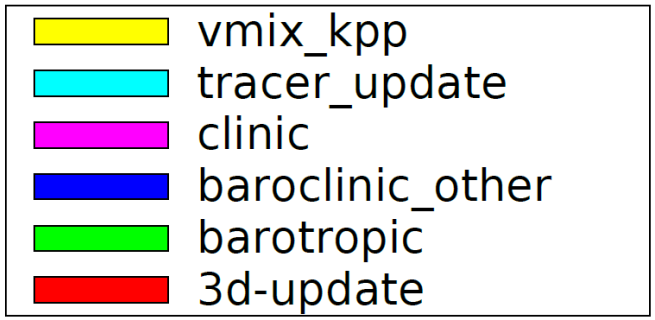
A performance profile of POP shows a complete lack of a **hot spot** (a single piece of code that most of the time)!

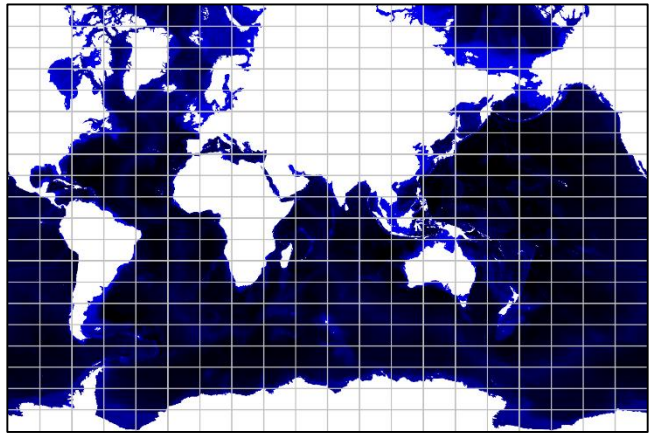
– So POP is **not** an obvious GPU candidate!

Source: Ben van Werkhoven, VU

Are ocean models suitable for GPUs?



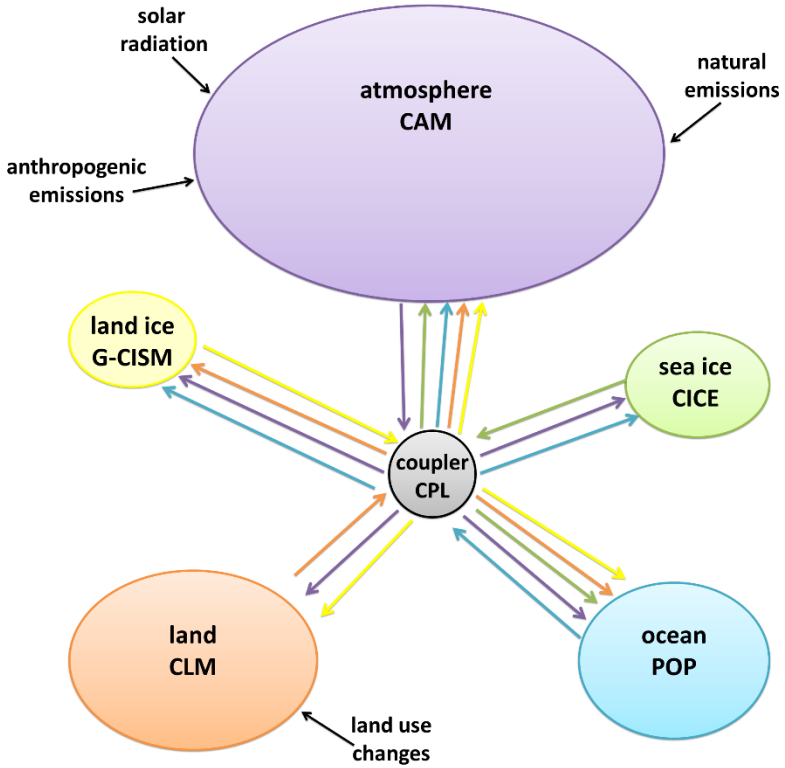




Can we run climate models distributed?

Using **distributed computing** we run different submodels on different supercomputers.

This allows us to **scale** each submodel to **higher resolutions** and combine different **architectures** (for example a traditional machine + a GPU machine).



Source: climatesight.org

Top 500 #10
2.9 PFlop/s

SUPERMUC (GER)

10G

Top 500 #184
223 TFlop/s

CARTESIUS (NLD)

GPUs

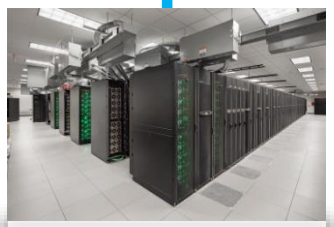
10G

114 TFlop/s

EMERALD (UK)

GPUs

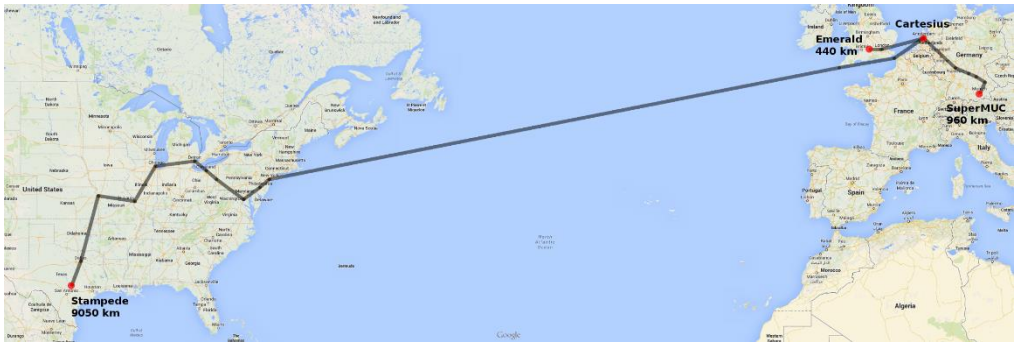
10G



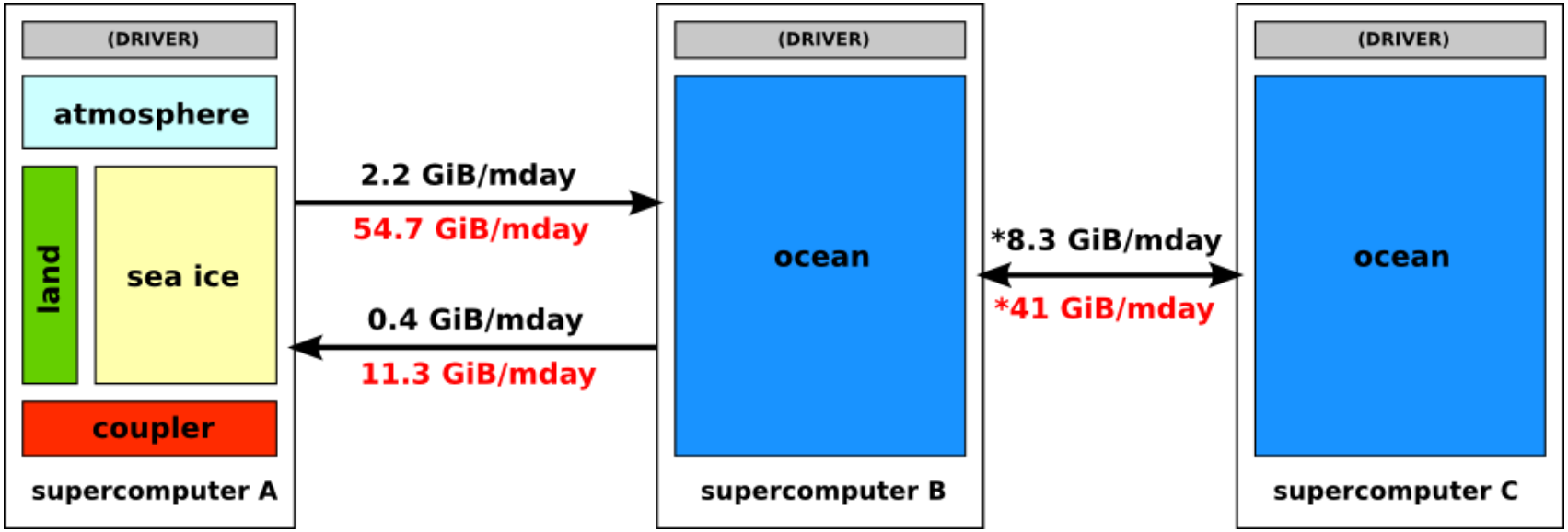
STAMPEDE (USA)

Top 500 #7
5.1 PFlop/s

Xeon Phi



Combine 4 supercomputers from EU and the US and 10G network links to run an Earth System Model at 0.02° (2x2 km) resolution.



Split ocean model using **geography aware load balancing**

(van Werkhoven et al, GMD 2014)

8.3 Gigabyte per exchange for 10 km resolution

41 Gigabyte per exchange for 2 km resolution

streaming traffic, latency sensitive!

Performance Models

This research also lead to the paper:

Performance Models for CPU-GPU data transfers, Ben van Werkhoven et al, CCGrid 2014 (best paper nominee)

Will be explained in detail later today in talk of Henri Bal.

Source: Performance Models for CPU – GPU Data Transfers, Ben van Werkhoven et al., 14th IEEE/ACM International Symposium on Cluster, Cloud and Grid Computing (CCGRID), best paper nominee, Chicago, IL, May 2014.

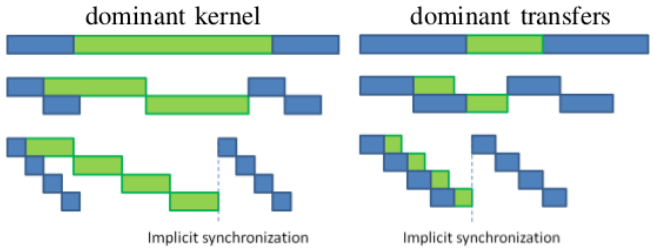


Fig. 1. Diagrams showing possible overlap for devices with implicit synchronization and 1 copy engine when using 1, 2, or 4 streams.

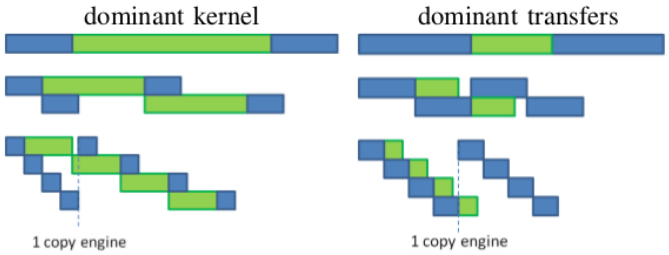


Fig. 2. Diagrams showing possible overlap for devices with no implicit synchronization and 1 copy engine when using 1, 2, or 4 streams.

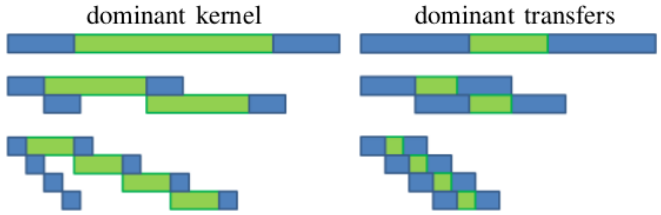
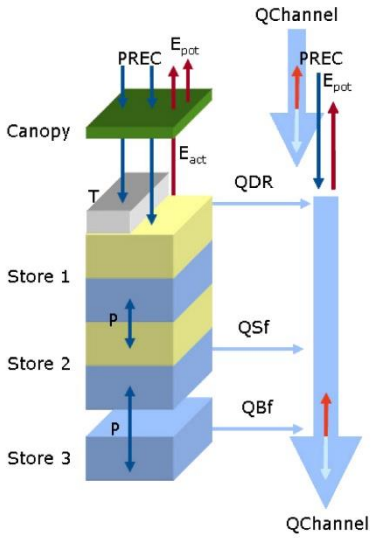
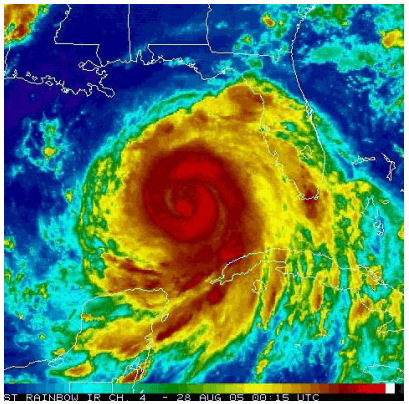


Fig. 3. Diagrams showing possible overlap for devices with no implicit synchronization and 2 copy engines when using 1, 2, or 4 streams.

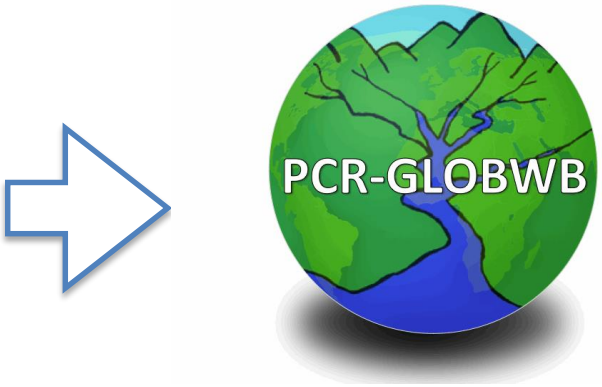
eWaterCycle: Global Hydrological Forecast System

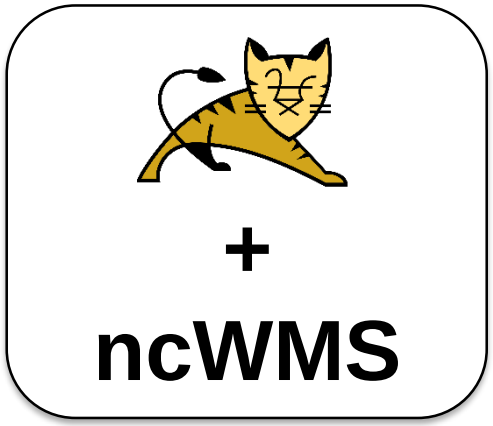
- **Global @ 10Km resolution groundwater and discharge forecasts**
- **GFS Ensemble meteorological forcing**
- **H-SAF Soil moisture observations**

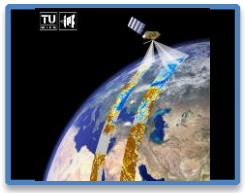





GFS
Global Forecast System


PCR-GLOBWB


+
ncWMS

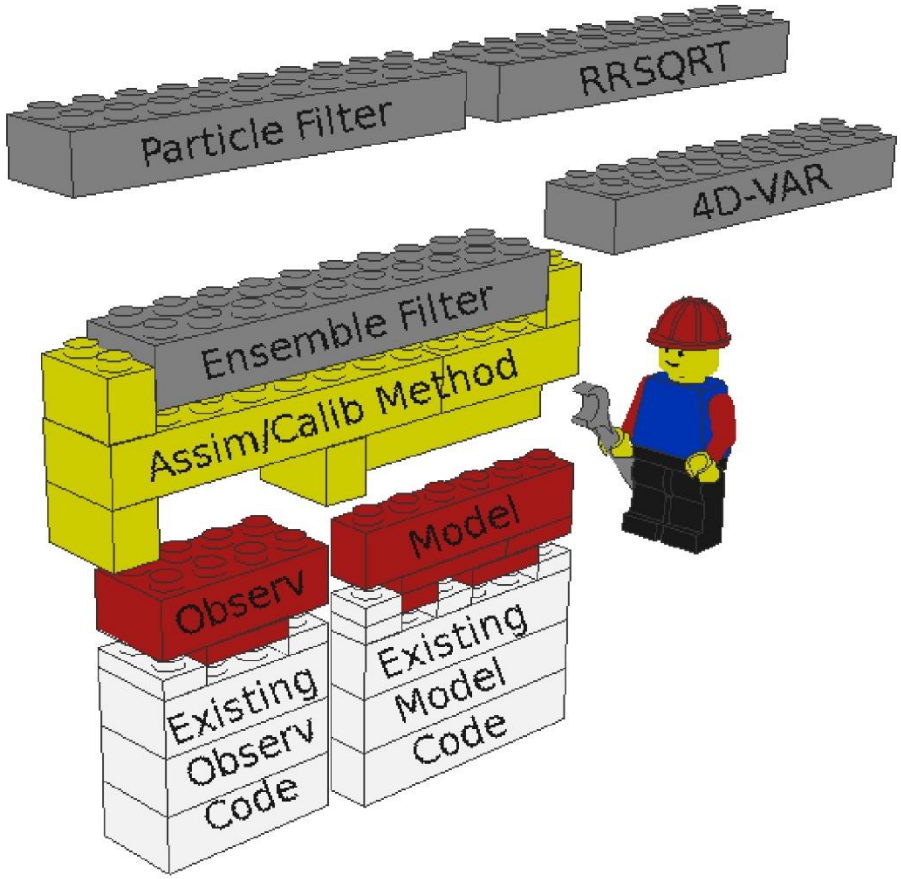



OpenDA


WaterCycle
TUDelft
netherlands eScience center

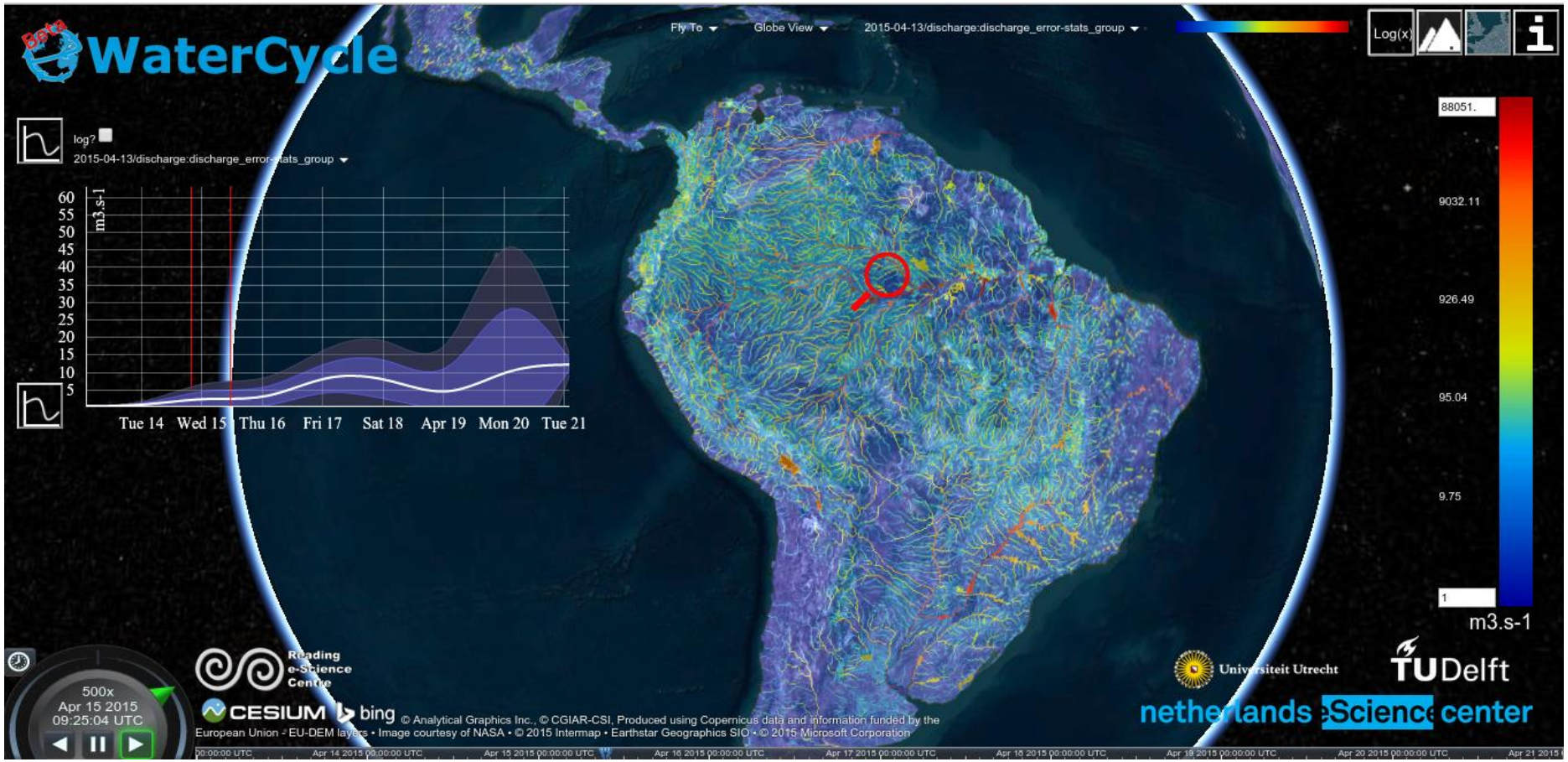
OpenDA

- **Open Source data-assimilation toolbox**
- **LGPL, Java**



http://www.openda.org

http://forecast.ewatercycle.org/

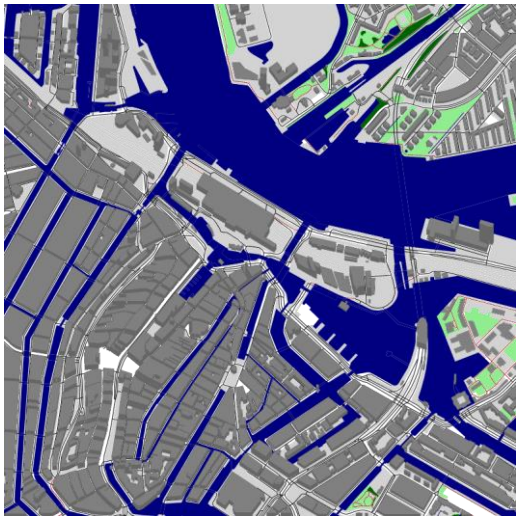


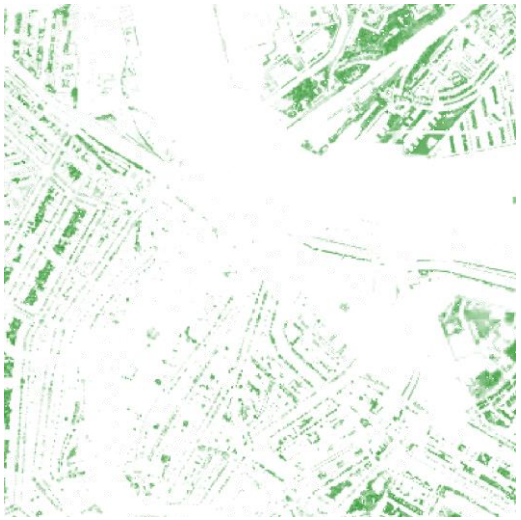
Reanalysis and forecast of urban heat

- **Automatic retrieval and conversion local city data (stations and bicycles)**
- **WRF high resolution urban hindcasts and forecasts**
- **Zip code level heat wave indices**

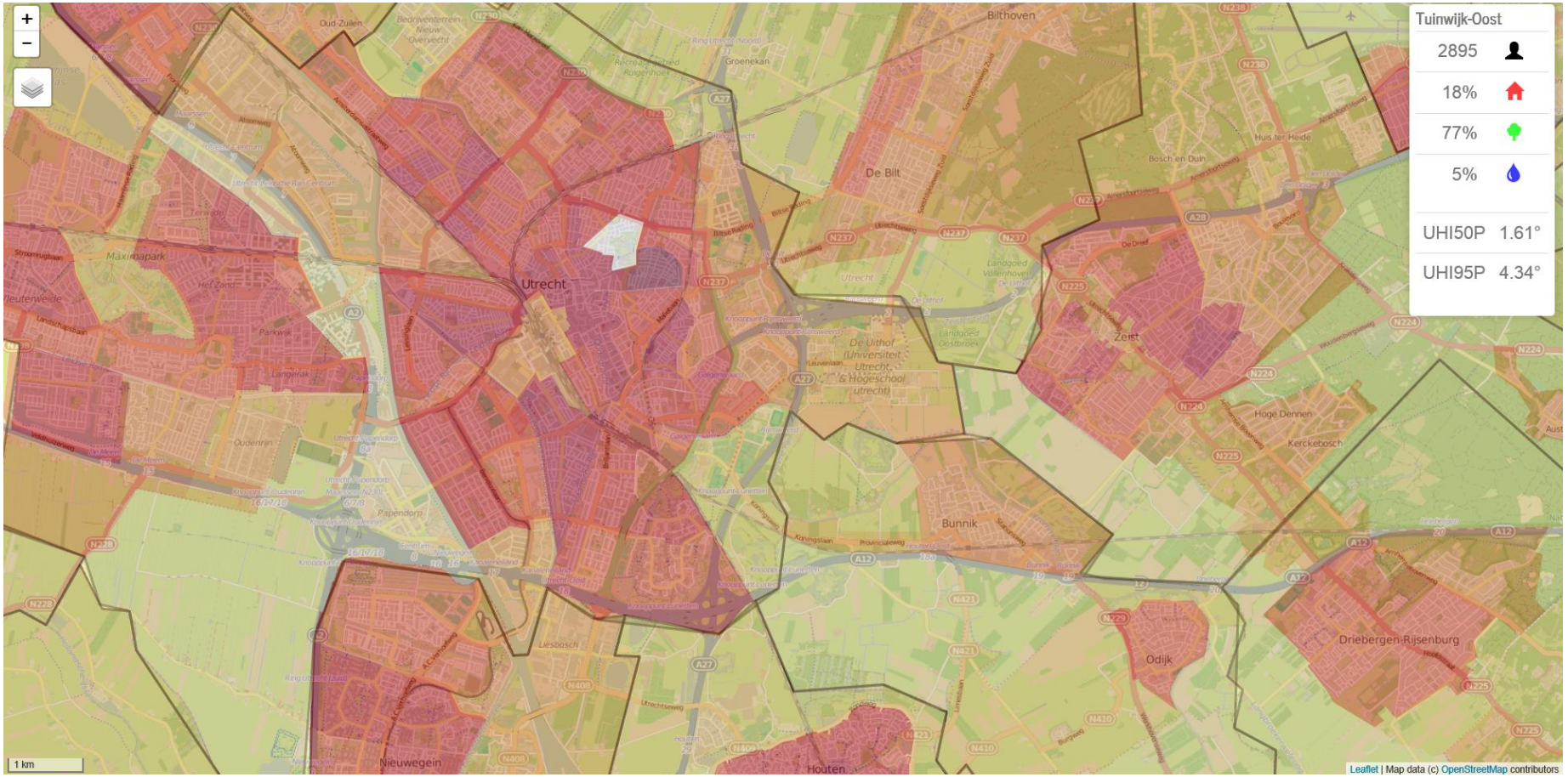
Local weather forecasting

- **WRF 3.5, 100 m horizontal resolution with urban parameterization (triple nesting)**
- **Selection of datasets: TOP10NL, satellite imagery, AH2 (height map), CBS data**
- **Data processed using PostGIS/Postgres and GDAL**





http://www.met.wau.nl/Summerinthecity/pages/fullscreeninfo.html



3rd National eScience Symposium

Amsterdam ArenA
8 October 2015

Accelerating Scientific Discovery Symposium

SCIENTIFIC TRACKS

Humanities & Social Sciences

Life Sciences & eHealth

Environment & Sustainability

Physics & Beyond

Computer & Data Sciences

KEYNOTE SPEAKERS

Prof. José van Dijck
KNAW

Prof. Leonard Smith
London School of Economics

Prof. Tony Hey
University of Washington

This event is free to attend, but registration is required:
www.eScienceCenter.nl

eScience center

in partnership with

NIKHEF

Royal Netherlands Meteorological Institute
Ministry of Infrastructure and the Environment

DTL

eHumanities
Royal Netherlands Academy of Arts and Sciences

IPN