NES TDCC "2022" Roadmap

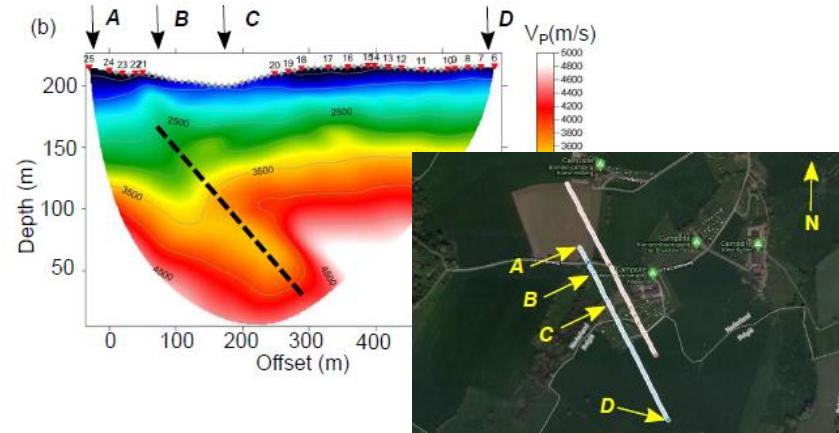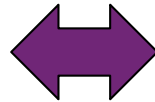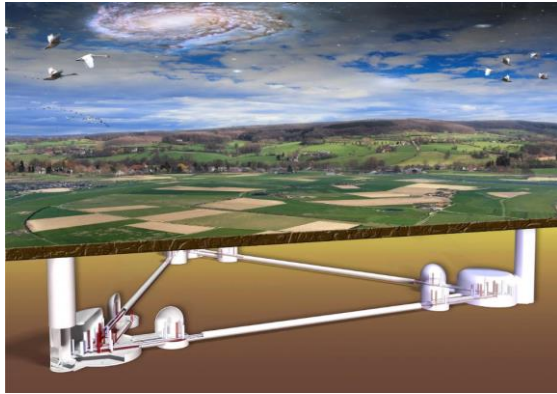# Starting the Natural & Engineering Sciences Thematic DCC

Nikhef Maastricht University

David Groep
September 2022

for the NES TDCC writing team
Anthony Brown, David Groep, Marta Teperek,
Irene Debussy, Vera Sarkol, Wim Som de Cerff

'NES': a heterogeneous domain with common challenges

Case study 1: Einstein Telescope seismic studies … ('E-TEST')

Data collected here is also useful for many others - outside of the ET planning - and we build on existing competences from domains decidedly 'non-GW' … that span many 'local' organisations (UM/Nikhef, KNMI, VU/Nikhef, UvA/Nikhef, Liege, Aachen, …)

Opportunities for collaboration

Case study 2: quantum research: quantum mechanics, physics, materials science, electrical engineering to come together with best practices

Quantum computing's reproducibility crisis: Majorana fermions

The controversy over Majorana particles is eroding confidence in the field. More accountability and openness are needed – from authors, reviewers and journal editors.

Sergey Frolov ✉

https://www.nature.com/articles/d41586-021-00954-8

Experiments to find Majorana signals are performed by loading a nanowire into a dilution refrigerator capable of cooling it down to close to absolute zero. Credit: HGA Architects and Engineers
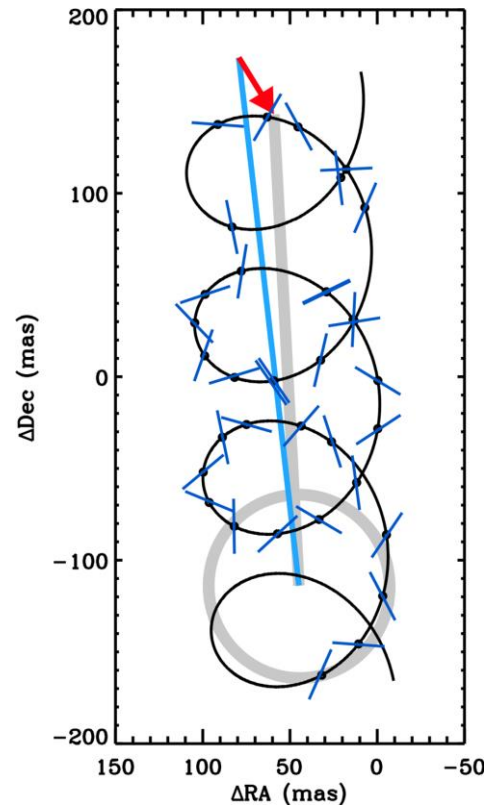
Starting the NES Thematic DCC
Slide and imagery: Marta Teperek, 4TU.RD and TUDelft

Nikhef

Opportunities for collaboration

Case study 3: Long term data and software archiving

> Data such as from Gaia are scientifically useful 'forever'; e.g. astronomical data from a century ago are being reprocessed and combined with Gaia data

- Long term (10s to 100s of years) preservation of raw and processed data
  - Including calibration data and models etc
- Preserve the software needed to handle/process/interpret the data
- Facilitate reprocessing of data in the future
  - E.g. new insights into forward modelling the data, combination with new/other data

See https://doi.org/10.1007/978-1-4614-3323-1_2 for more details on the Gaia example

Exoplanet detection by combination Hipparcos and Gaia data (spanning 25 years)
https://doi.org/10.1038/s41550-018-0561-6

Slide and imagery: Anthony Brown, Sterrewacht Leiden

The challenge for the NES TDCC

- Connect **research communities** to **existing infrastructures**, tools, services and expertise

- Foster a **coherent ecosystem** of digital services, data and software across the domain.

- Stimulate collaboration and exchange of **best practices across disciplines**

- Increase **expertise among researchers** so they can steer more initiatives in their own field

- Provide continuous feedback to NWO on **research support needs**
  to inform NWO's (future) resources distribution on digital competence topics

- Facilitate connections also **between thematic DCCs** for cross-disciplinary research

Luckily, the NES TDCC stands not alone

- **Local DCCs** are essential to identify, connect and support the domain-specific needs, and the TDCC can help to extend beyond institutional, and national, borders

- there's a rich ecosystem on infrastructure, software, data and expert groups surrounds us, with **SURF**, the **Netherlands eScience Center**, **LCRDM**, NWO's Open Science Fund, *&c*

- there is inherent, structural, expertise in the **Roadmap Facilities of the PC-GWI** (you!)

- strong international collaborations & forums (RDA, IVOA, ECFA, …) and the ESFRIs

- we now have our new Open Science 'Regieorgaan' coordination body

Initial challenges within the NES domain

- **Fragmented landscape** – insufficient connections
  between research communities, existing infrastructures, tools, services & experts

- Lack of interoperability of data and of workflows: **especially the 'I' and 'R'**
  & need to integrate workflows to put FAIR into practice and bring data to 'life'

- Insufficient **training capacity** on digital competences

- Insufficient support long-term preservation & sustained maintenance of **software**

- Insufficient computational **reproducibility** practices
  (e.g. implementation of provenance standards and agreement on level of detail)

- Lack of clarity and guidance about **dealing with IP** and legal issues

- Insufficient **collaboration between disciplines** across the domain
  which could facilitate the adoption of FAIR practices

The NES Roadmap v1 is now ready!

Based on the extensive consultation process

- PC-GWI, science deans, LDCCs,
  Netherlands eScience Center, SURF, LCRDM, and you:
  the national Roadmap Facility digital expert group, …

- stakeholders* welcome
  to **shape programme and projects**

- **anyone can use the output**, join the
  trainings, and use the network!

Roadmap Thematic Digital Competence Center - v1.0

Domain Natural and Engineering Sciences

Lead authors

Anthony Brown, Universiteit Leiden
David Groep, Nikhef
Marta Teperek, 4TU.ResearchData

Contributing authors

Irene Dedoussi, TUD
Vera Sarkol, CWI
Wim Som de Cerff, KNMI

NWO-secretaries

Erik van Aert, Domain Sciences
Paul Blank, Domain Applied and Engineering Sciences

Consultations during the Roadmap preparation

During the process of Roadmap drafting, the writing team, as well as NWO, took a proactive stance for seeking input from the broader community. Input from members of the 4TU.Federation (technical sciences) was gathered regularly and at various levels (university rectors, faculty deans and research directors, library directors, individual researchers and data stewards). Feedback from the astronomers and KNMI was ensured through direct representation in the writing team. There were also several consultations organised with the NWO institutes.

In addition, to reach out to all the research performing organisations with natural or engineering sciences faculties, a dedicated consultation was organised to which representatives from all local DCCs were invited. During this session, the writing team presented the vision for the Roadmap and received feedback from the participants, which directly shaped the Roadmap.

Given that the thematic DCCs are network organisations, they need to rely on infrastructure provided elsewhere. Therefore, the writing team had several bilateral conversations with colleagues from SURF. In addition, members of the writing team spoke on several occasions with colleagues from the eScience Center to

*) contribution can also be in-kind with your time investment!

Bottleneck 'projects'

- Community building
  - the national expertise map and thematic survey

- Creating a Training Hub
  - leveraging existing communities and material

- Facilitating long-term software preservation
  - curation guidance, best practice licensing and publishing, models for sustaining research software

- Support for metadata standards and vocabularies for FAIR data
  - linking with RDA, promoting community standards to new users

Beyond the bottleneck 'projects'

- these will be complemented with more & targeted projects in years following
  - identified by a domain-wide continuous consultation process
    in the NES TDCC governance board

- support together a coherent work programme and foster collaboration
  - by centrally defining that work programme, and calling for teams to work
    together on realizing the NES work programme

- project calls will be to execute a specific part of the work programme
  - *in a collaborative way*
  - call process will be promoting collaboration

... the NES TDCC looks for your input and appreciates your participation!

all thanks should go to the writing team!
and any blame: please direct that to me

"YOUR TDCC NEEDS YOU"

The NES TDCC writing team:
Anthony Brown, David Groep, Marta Teperek,
Irene Debussy, Vera Sarkol, Wim Som de Cerff
NWO support by Erik van Aert and Paul Blank
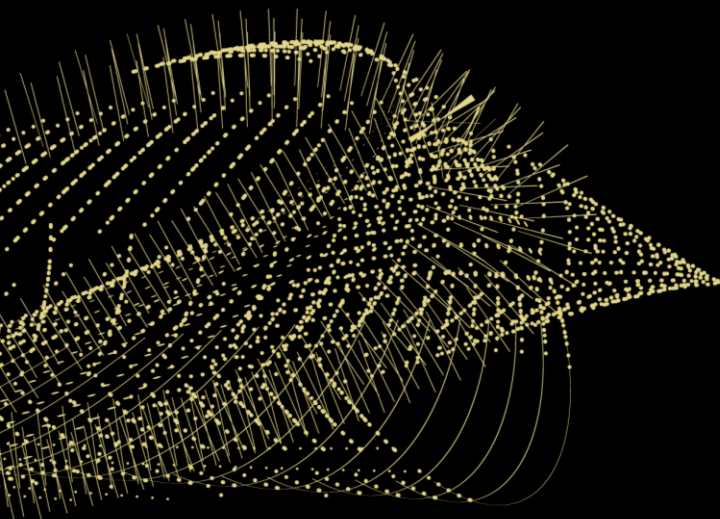
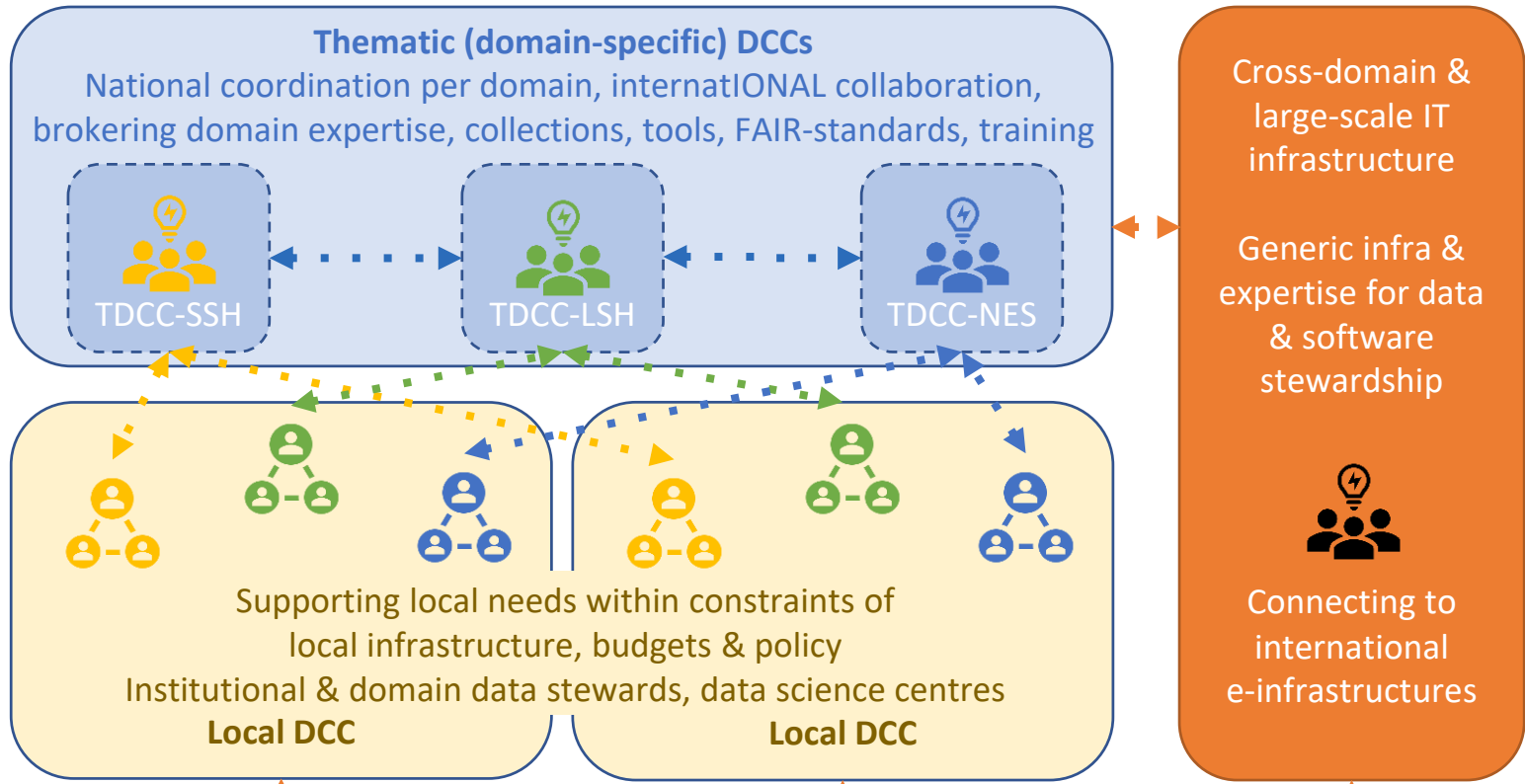Maastricht University Nikhef

David Groep
davidg@nikhef.nl
https://www.nikhef.nl/~davidg/presentations/
https://orcid.org/0000-0003-1026-6606

Supplementary material

Three domain-specific digital competence centres

Thematic (domain-specific) DCCs
National coordination per domain, internatIONAL collaboration, brokering domain expertise, collections, tools, FAIR-standards, training

TDCC-SSH

TDCC-LSH

TDCC-NES

Cross-domain & large-scale IT infrastructure

Generic infra & expertise for data & software stewardship

Connecting to international e-infrastructures

Supporting local needs within constraints of local infrastructure, budgets & policy
Institutional & domain data stewards, data science centres
Local DCC

Local DCC

Starting the NES Thematic DCC
Image by: Ruben Kok, LSH TDCC and DTL

Nik|hef

Building the NES TDCC – governance

Governance Board

*composed of contributing **members from all disciplines***

- Develops the consortium agreement
- Determines the priorities and **work programme**
- Ensures alignment with national digital research priorities and other thematic domains
- Guides the evaluation of project proposals

Coordination Team

Supports the Board and implements the programme

Network Manager

Community Coordinator

Starting the NES Thematic DCC
Image by Marta Teperek, 4TU.RD and TUDelft

NES TDCC organisation: coordination team

COORDINATION TEAM

NETWORK MANAGER (CONSORTIUM COORDINATOR)

- Represents the consortium
- Drives governance and membership discussions
- Structures the project portfolio and ensures alignment with the programme
- Liaises with infrastructure, tool and resource providers
- Oversees communication and online presence

COMMUNITY COORDINATOR

- Identifies and liaises with community members
- Identifies community needs
- Maps the infrastructure, existing resources and tools
- Connects people from different disciplines and connecting them with infrastructure, resources and tools

Starting the NES Thematic DCC
Image by Marta Teperek, 4TU.RD and TUDelft