14th meeting of the DTL Partner Advisory Committee

May 19, 2021

Participants

Name/Affiliation
Jeroen de Ridder, UMCU & Oncode Institute
Alain van Gool, Radboudumc (and X-omics, EATRIS)
Peter van Dijken, TNO
Dick de Ridder, Wageningen University
Ingrid Dillo, DANS
Cees Hof, DANS
Patrick Kemmeren (Prinses Maxima Centrum)
Barend Mons (LUMC, GO FAIR)
Henk-Jan van den Ham, ENPICOM B.V.
Filip Pattyn, ONTOFORCE & UGent
Andry Klumpert (Vancis)
Lars Ridder, eScience Center
Michel Dumontier, Maastricht University
Kees de Gooijer
Solon Pissis, CWI
Jantine Dirksen (Quaero Systems)
Renger Jellema, DSM
Joris van Eijnatten, eScience Center
Ronald van Schijndel (Amsterdam UMC; starting from 15:50)

Peter Hilbers, TU/e, chair
Maurice Bouwhuis, SURF
Eric de Vries, Digeketen
Eva van Ingen, Elevate Health
Han Bakker, LabServant
Harry van Haaften, The Hyve
Jaap Heringa, VU & ELIXIR-NL
Jan-Willem Boiten, Lygature & Health-RI
Joep van Wijk, NWO
Martin den Heijer, AmsterdamUMC
Mijke Jetten, DTL
Robert Veen, Erasmus MC
Ronald van Schijndel, Amsterdam UMC Locatie VUMC
Solon Pissis, CWI
Dick de Ridder, Wageningen UR
Marielle Gallegos Ruiz (Roche)
Maurice Bouwhuis (SURF)
Hans Niendieker, Ivido
Peter van Dijken - TNO

DTL team: Ruben Kok, Mijke Jetten, Rob Hooft, Merlijn van Rijswijk, Mascha Jansen

Useful information:

The slides from the meeting can be found here.
<u>Annexes:</u> https://www.dtls.nl/about/organisation/pac/documents-pac-meetings/

The report of the 13th meeting has been approved without modifications

Table of Contents

1. Presentation: *Open science in medical evidence generation: the journey of EHDEN,* Kees van Bochove - The Hyve

Kees van Bochove gave a presentation on the Open Science approach of EHDEN. EHDEN (*European Health Data & Evidence Network*) is a European project of the Innovative Medicines Programme (IMI), aimed at developing a European federated health infrastructure, based upon open source software.

Essential to EHDEN is its training programme. EHDEN also has a harmonisation fund (mapping to the OMOP common data model), currently in its 4th round.
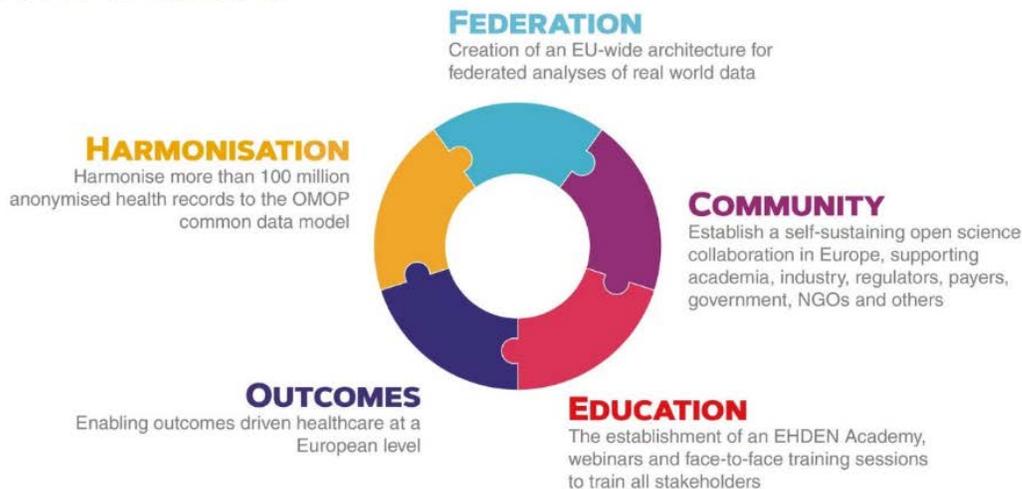
Now available for all: EHDEN project deliverable D4.5: Roadmap for Standardisation Solutions, comparison of standards to the OMOP Common data model (OMOP: *Observational Medical*

Outcomes Partnership). Kees also spoke about [OHDSI (Observational Health Data Sciences & Informatics)](#), which focuses on standardizing vocabularies and mapping between ontologies. OHDSI also aims at standardizing analytics, resulting in a change towards global systematic analysis of (health) data.

Example of Covid-19 data - EHDEN organised a Covid-19 study-A-thon March 2020: take a medical research question, organise an event to answer that question (instead of months). In March, EHDEN focused on twelve questions, for instance on effects of hydroxychloroquine. Within weeks, the study-a-thon resulted in tangible output, including publishing the data and paper preprint for all 12 questions.

Main takeaway: medical evidence generation is changing through open science, including open data.

Questions/remarks/ answers raised during the presentation:

- Question: "Did you use tools, or was it done by hand?"
  Answer: "Map data to OMOP so it can be used, we have tools that help with mapping. Normally, you have to go manually to the data documentation to map, we help to make that process easier via the tool. We should encourage people to make data more FAIR at the source."

- Question: "Could you further explain the 'harmonisation fund' approach?"
  Answer: "Any organization that has custody of data and is in the EU in a broad sense can apply for money. To be spent on staff or hire a company (encouraged). We need a distributed network of experts, we can't leave that to one organisation. Kick start this network, with paid efforts."

- Remark: "Sounds like the hotel approach so you have enough capacity and expertise. Please further discuss with us to see how we can together continue and stimulate this.

- Question: "Does one have to be a EHDEN or OHDSI partner to access the data and is it entirely done by data visiting or only when you do some sort of studyathon?" Answer: "Data partner onboarding is done via our calls. In studies we are in early days. Anyone with OMOP data can participate in EHDEN. The PHT concept could be with virtual machines a major step forward, which would be great for Health-RI."

- Remark: "Connection between Health-RI and EHDEN needs to be made, but also the H2O project (also using OMOP) is interesting. Kees is open to discuss this."

- Remark: "A quick analysis of FAIR principles vs OMOP is also included in https://zenodo.org/record/4474373"

- Remark: "Let's discuss these options to link EHDEN and FAIR implementation approach at Health-RI, also in the framework of mapping to FHIR."

- Remark: "Starting with a focus group on FHIR/FAIR/OMOP and super-conceptual models. I would certainly participate. Also critically important for TWOC, health-RI and NPOS."

2. Presentation: *FAIRsFAIR - Fostering FAIR Data Practices in Europe, Ingrid Dillo, KNAW*

Ingrid Dillo introduced the FAIRsFAIR project. This H2020 project is funded by the EU, and led by DANS. Other Dutch partners include UvA, SURF and DTL. The objective of FAIRsFAIR is to offer generic practical solutions for FAIRification, throughout the data life cycle. Various results are available via Zenodo: policy recommendations for publishers and funders, guidelines, certifications and support,

developing tools for FAIR data assessment, on training and professionalisation of FAIR, run data stewardship training via RDA/CODATA schools.
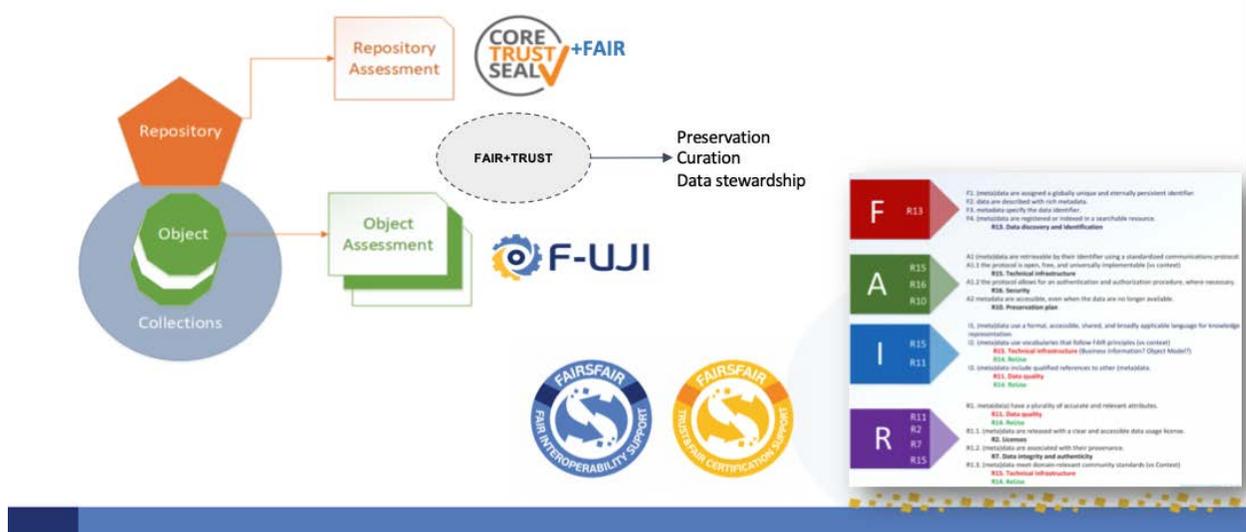
The project focuses on three areas:

1. policy and practices,
2. certification,
3. training and professionalization.

TRUST principles for repositories were developed in the framework of RDA (international and community-based) and CODATA; Transparency, Responsibility, User Focus, Sustainability, Technology (TRUST).

Repositories play a role in basic FAIRness & keeping data FAIR over time. Therefore FAIRsFAIR focuses on FAIR assessment of these repositories: CORE Trust Seal combined with assessing FAIRnes. Mapping available with FAIR, including FAIR maturity model. Furthermore, the FAIRsFAIR project developed FAIR data object assessment metrics, the F-UJI tool for automated machine actionable FAIR Data assessment, and a web-based graphical user interface. FAIR is not an absolute qualifier. Ingrid gave an example of a data set that met the FAIR criteria including a score for the FAIRness of the dataset: level of

FAIRness as percentage. Green metrics (criteria that was met), red metrics (criteria that were not yet met).

<table>
<tr><td bgcolor="magenta">Questions/remarks/ answers raised during the presentation:</td></tr>
</table>

- Question: "How much of the FAIRsFAIR criteria can be used for making data FAIR at the source?"
  Answer: "Focus at the full data life cycle so that part is included. Tool for that as well (FAIRAWARE tool: https://fairaware.dans.knaw.nl/) - that helps researchers to think about the elements that are relevant from the start of the project. Plus: importance of training for current and future researchers, in educational programs."

- Question: "Are there metrics to test the TRUST principles? And how does the tool's operation/metrics compare to the FAIR Evaluator by Wilkinson et al?"
  Answer: "We are not yet there."

- Question: "How does the FAIRsFAIR tool compare to the 'FAIR evaluation Services' tool by Mark Wilkinson?"
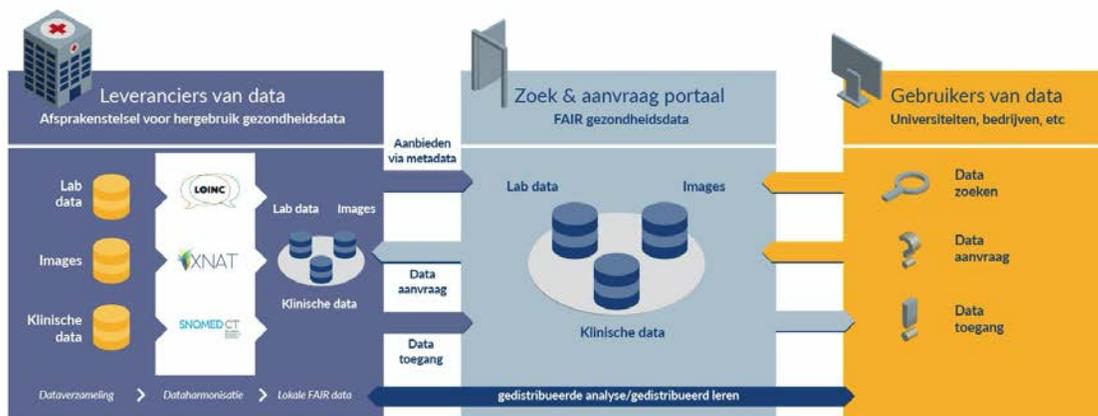 Answer: "F-UJI tool is relatively simple, therefore there is a broad uptake, also by NIH in the USA."

- Remark: "The FAIRsFAIR tool is looking only at the FAIRness of data in repositories, and is not taking all aspects into consideration that Mark Wilkinson's approach does."

- Remark: "We need both (tools) as Mark is really focussing on 'computers testing machine actionability of the data in 'independence' and is not for 'common users, like F-UJI. In the WELLCOME TRUST project I think both will be explored/exploited. Also good to show no provider lock in :)"

3. Update by Ruben Kok - *Development of Health-RI as national data infrastructure for health research & innovation*

Ruben summarized recent developments from within and outside DTL's network, with a specific focus on the development of Health-RI, which has recently been awarded substantial funding from the Dutch Government.


Recent developments
- DTL has contributed in realizing a national COVID-19 observational data portal in Health-RI, in close collaboration with ZonMw and GO-FAIR. (https://www.health-ri.nl/covid-19-data-portal)
- ELIXIR:
  - ELIXIR Converge project has launched an RDM-kit for life sciences research with very practical guidelines and tools for research data management in the life sciences: https://rdmkit.elixir-europe.org/.
  - ELIXIR-NL has contributed, amongst others the Data Stewardship wizard.

- ○ ELIXIR webinar on professionalising data stewardship: recording (https://elixir-europe.org/events/elixir-webinar-towards-professionalising-data-stewardship) and slides (https://doi.org/10.5281/zenodo.4769514) are available.
- National Programme Open Science (NPOS) Project report on professionalising data stewardship: https://doi.org/10.5281/zenodo.4320504. The above webinar slide deck includes a summary of this report as well.
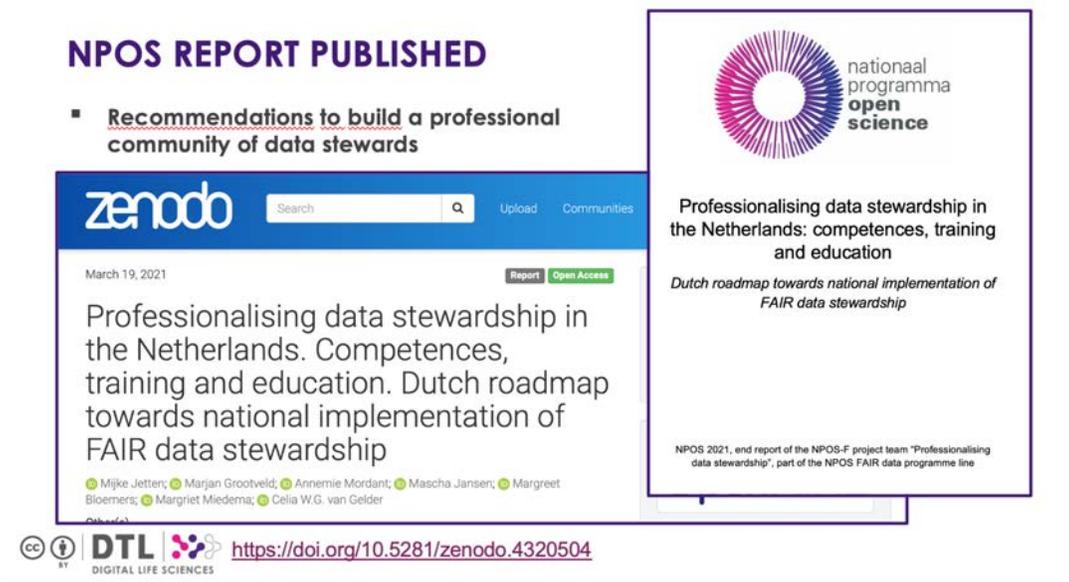
Health-RI grant National Growth Fund

Health-RI has been awarded a major grant of 69M€ to develop the national data infrastructure for health research and innovation. The goal is to make health data available for research and innovation purposes, by public and private stakeholders, taking into consideration the legal and ethical issues around privacy-sensitive data. FAIR implementation will be key to Health-RI.

The Health-RI infrastructure model will be based upon the creation of a national trust framework and data governance model around health(-care) data, involving the government and the broader field of stakeholders. Currently we are preparing to set up a network of regional Health-RI nodes, around the University Medical Centres and the Eindhoven region, connecting other partners in the field, such as hospitals, universities, institutes and also companies. Health-RI will effectively form a multi-community organisation to design and build the national health data infrastructure. The FAIR-based model of Health-RI will build on the experiences with the Covid-19 data portal.

How we work: action lines of Health-RI

1 National trust framework
A. Citizen control
B. Interoperability & linking standards
C. Data governance policy & agreements
D. Public-private collaboration models

2 FAIR health data
A. Establishing hub and nodes model with regional nodes
B. Establishing central hub and international connection
C. Facilitating federated analysis and federated learning

3 One-stop shop
A. Central FAIR data catalogue
B. Developing a portal for data driven research and innovation
C. Analysis software and digital workspace
D. Support, training & dissemination

0 Management & organization
A. Stakeholder management
B. Sustainable business model
C. Communication

- Until 2020: Set up operations (financed by NFU, Health Holland, ZonMw)

- 2020-2023: Founding and financing (NFU) to prepare a health data infrastructure

- 2021 onwards: Health data infrastructure / Groeifonds funding!

health RI enabling data driven health

Questions/remarks/ answers raised during the presentation:

- Remark: "Would be good to forward this webinar [on data stewardship professionalising data stewardship] to EATRIS and BBMRI communities or having it again with groups from those networks.?"
Answer: "Details (links) are available via the notes document, and we will indeed spread the information! Please do this to Eliis of EATRIS-Plus."

4. Presentation: *Development of a layer of thematic Digital Competence Centres in the Dutch data landscape, Joep van Wijk - NWO*

Joep van Wijk presented the plan of NWO to develop so-called *Thematic Digital Competence Centres* (TDCCs). This instrument is a follow-up of recommendations by the ICT committee Apers to invest in digital infrastructure (HPC, eScience Centre, local DCCs, thematic DCCs).

Thematic DCCs are foreseen as overarching 'interuniversity' digital competence centres (overall budget 2.5 mln/yr). The funding goes to setting up the thematic DCCs (organising the network), but also running projects with local DCCs from universities and UMCs. Thematic DCCs should be concentrated focus initiatives, and the result of cooperation and self-organisation in the field.

There will be three focus areas for thematic DCCs: Social Sciences and Humanities, Beta/Technical sciences, Life Sciences & Health. DTL will be invited as one of the partners in the life sciences and health domain, led by Ruben Kok.

The process entails three phases:

1) before summer finalise request to those three bridgebuilders,
2) drafting plan with the community,
3) fund the plan (expected late 2021).

The Thematic DCC's will be (co-)funded by NWO for 10 years, with evaluation after 5 years (go/no go). First 5 years NWO will on average fund 70%, after 5 years this will be reduced to 40%, additional co-funding necessary by the community. After 10 years, TDCCs must be sustained by the field.

NWO will give assignments to two persons (researcher and a facilitating partner) to set up a plan for a TDCC, starting with description from the (data) landscape, leading to a business case/plan.

Questions/remarks/ answers raised during the presentation:

- Remark: "NWO could support co-funding of this by ZonMW/VWS/EZK etc."

---

5. Presentation: *Exposome-NL – Integrating multifactorial data to study environmental drivers of health, Roel Vermeulen - Utrecht University*

Roel Vermeulen presented the Exposome NL project, which has received 17,4 million euro of NWO funding. It focuses on understanding the influence of environmental, social and/or lifestyle factors on

health and disease development. Using the power of big data, exposome research tries to link such exposures during the lifespan of a person to health and disease.

- Multitude of different data and data sources needed to get a clear picture on the exposure. Health endpoints measured in the epigenome, transcriptome, proteome, metabolome and microbiome.
- Important technology push in metabolomics to measure the chemical space in collaboration with Thomas Hankemeier from Leiden University.
- [Exposome Data Infrastructure (NIFER)](#)

Health Data + External Exposome + Internal Exposome: all data and meta-data needs to be made FAIR. This Exposome Data Infrastructure fits very nicely with Health-RI, also a dashboard is being provided.

Furthermore, the [Horizon Europe Partnership for Chemical Research Assessment](#) is being developed: a 200 million programme, with a large work package on FAIR data led by Roel Vermeulen. Problem: everyone sees that this is important, but partners are often not yet ready to fund FAIRification of data.

The project also links with the [Virtual Human Platform](#) for safety project, recently funded from the National Science Agenda (Nationale WetenschapsAgenda, NWA).