



The presentation will start soon



This webinar will be recorded

Recordings can be made available for knowledge dissemination.



Questions or remarks?

Please raise your hand by clicking on the “raise your hand”-symbol or use the chat.



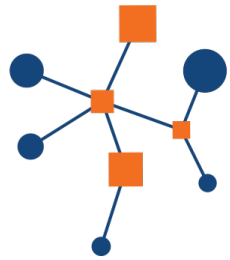
Mute your microphone

Please mute your microphone to minimize background noise by clicking on the “mute”-symbol on the top of your screen.



What can we improve for the future?

We are happy to hear any additional questions or feedback for our next expert meeting. Please contact us by mail: robin.verjans@lygature.org



BBMRI.nl
Health-RI partner



health RI
enabling data driven health

Expert Meeting Series 2021

The Next Step in Making Health Data Available for Research

Expert Meeting #1

09/09/2021



health **RI**

enabling data driven health

Wiro Niessen – CTO Health-RI

Erasmus MC, University Medical Center Rotterdam
Delft University of Technology



Honorering Groeifonds Voorstel

Realisatie lerend zorgsysteem via drie actielijnen in een hub & nodes model

1

Afsprakenstelsel

- A. Zeggenschap burger/patiënt
- B. Interoperabiliteit & koppelstandaarden
- C. Data governance processen & overeenkomsten
- D. Publiek-private samenwerkingsmodellen

2

FAIR gezondheidsdata

- A. Inrichting regionale knooppunten
- B. Inrichting centrale hub en internationale aansluiting
- C. Faciliteren gefedereerde analyse en gefedereerd leren

3

Een loket

- A. Centrale FAIR data catalogus
- B. Ontwikkelen one-stop shop voor data-gedreven onderzoek en innovatie
- C. Analysesoftware & digitale werkomgeving
- D. Support, training & disseminatie

0

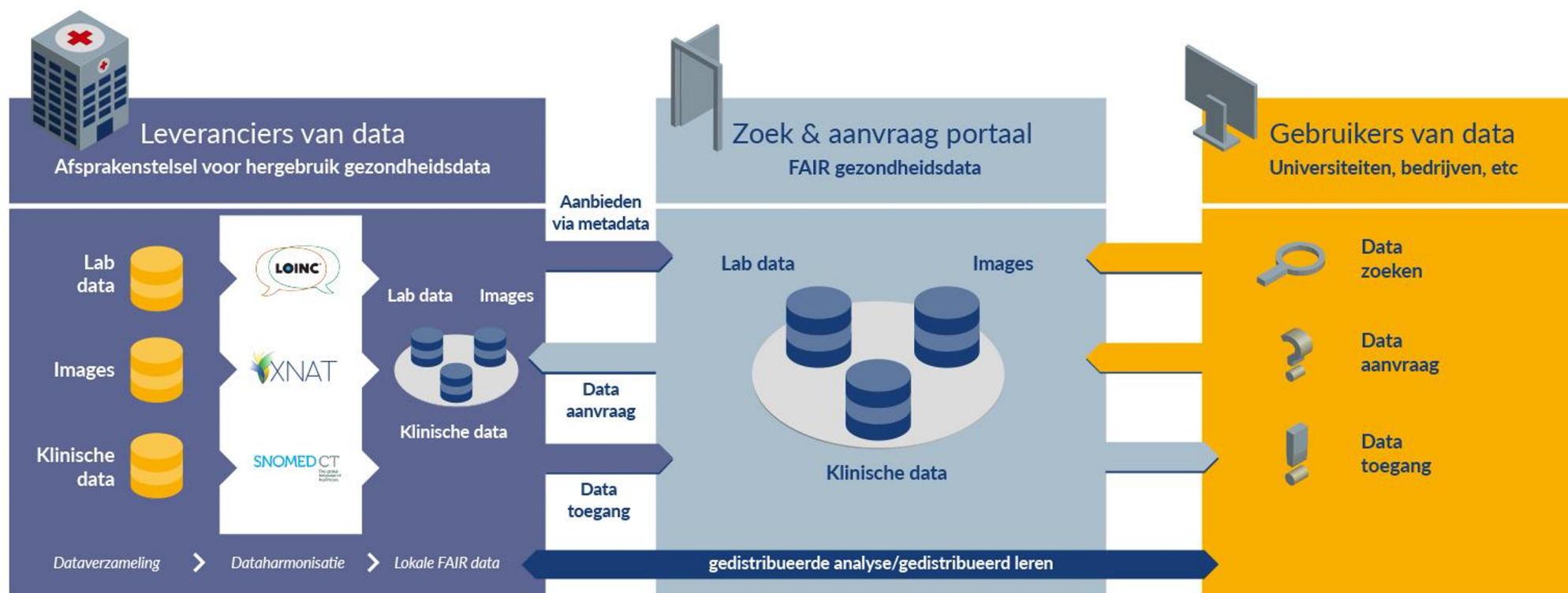
Management & organisatie

Communicatie, Management, Financiën



Building on COVID-19 use case

Towards a national COVID-19 observational data portal

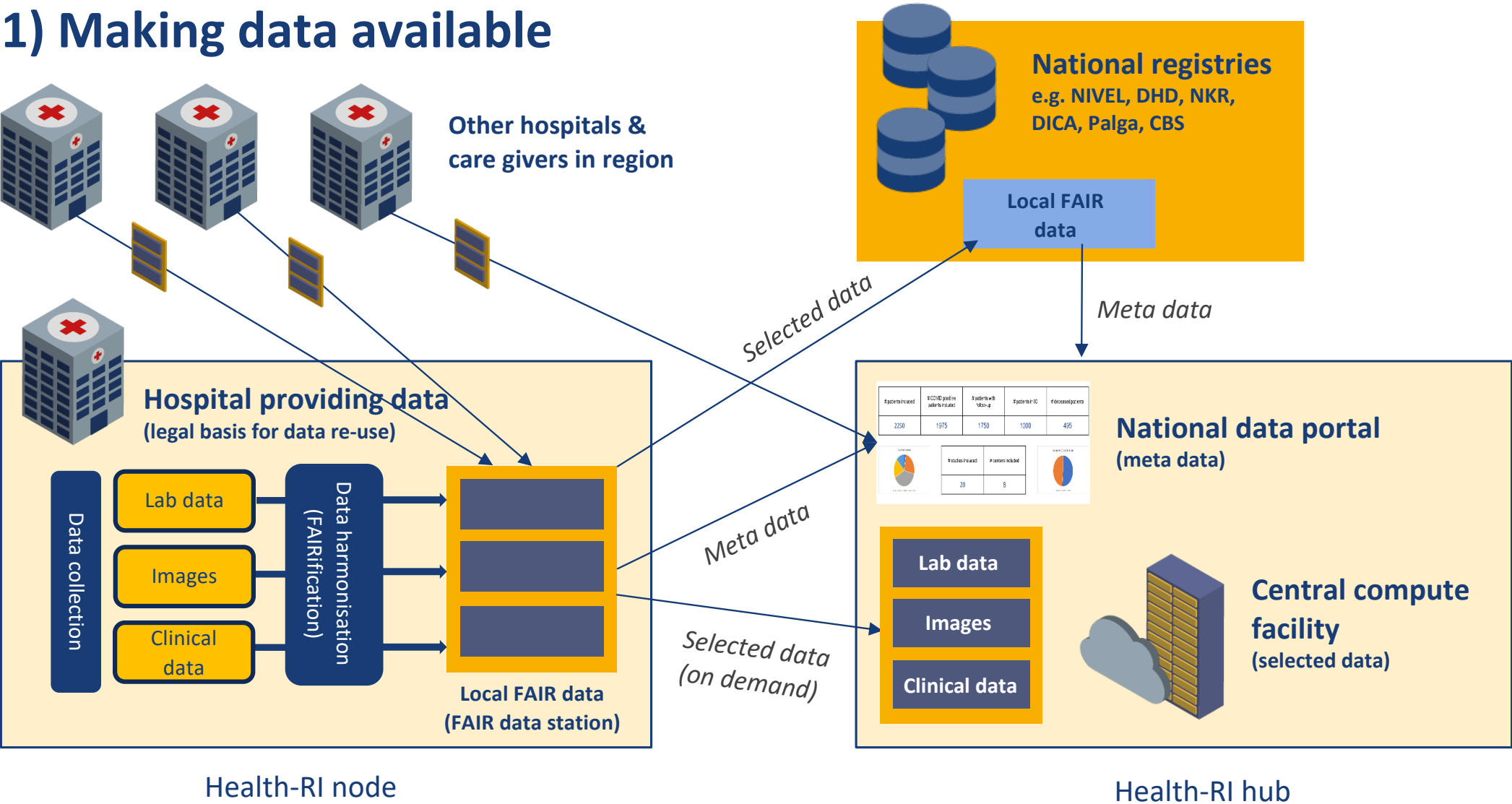


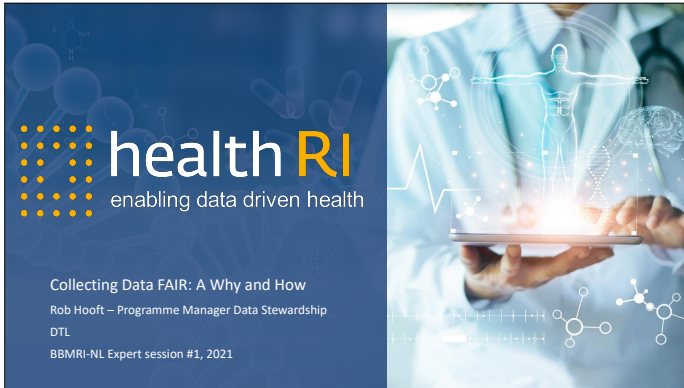
Hub&nodes model: principles for implementation

- Implementation of three overarching workflows has priority:
 1. Make FAIR data available at regional nodes
 2. Data discovery, access & governance
 3. Data analysis & learning (federated (PHT) & via workspaces (anDREa))
- Use cases should provide demonstrators for the practical usability of the implemented workflows and guide their development
- All activities should either
 - Support development of the regional node
 - Support development of the hub
 - Support the federated infrastructure connecting hub, nodes and other infrastructure



1) Making data available





Reusable

We willen graag dat de gegevens die we verzamelen in cohort en populatiestudies veel worden hergebruikt
We willen die data daarom graag zo bruikbaar mogelijk maken.

FAIR

In dit verhaal spreek ik over FAIR data, en de R uit FAIR staat precies voor Re-usable.

machines: FAIR is voor mensen en vooral voor machines: analyses opschalen vereist het automatiseren van processen en het minimaliseren van handwerk.

historie: In 2014 zijn op een meeting in NL de FAIR principes uitgedacht: wetenschappelijke data moeten F,A,I,R zijn; voor mensen en vooral machines!

Ladder: De principes zijn in principe onafhankelijk, maar toch vormen ze een soort ladder: Eerst maken we data F, daarna...
Als we naar herbruikbaarheid streven, zijn alle 4 de letters essentieel

Why?

Concrete suggesties hoe: Daarna zal ik voor elk van de letters van FAIR een paar suggesties geven hoe dit vorm kan krijgen

I would like to exploit common genotype-phenotype relations between Alzheimer's Disease and Huntington's Disease...
I need to combine AD and HD data...

I can help with that!

Source: Marco Roos

healthRI
enabling data driven health

5

Data zijn niet zonder meer koppelbaar.
Van tevoren dezelfde standaard kiezen helpt, maar is niet 100% noodzakelijk.

Er is dus voor hergebruik meer nodig dan een technische standaard: er moet vooral heel veel metadata beschikbaar zijn. De data moeten worden beschreven. Wat zijn het voor gegevens, hoe zijn die er gekomen, hoe zijn ze gecodeerd en misschien zelfs waarom.

DOES NOT COMPUTE

Here's my data, have fun!

Here's my data, have fun!

Here's my data, have fun!

Source: Marco Roos

health RI
enabling data driven health

Source: Marco Roos

healthRI
enabling data driven health

We kunnen die FAIR-interoperabiliteit niet makkelijk achteraf toevoegen. Beter is het om daar zo vroeg mogelijk mee te beginnen: daar waar de data wordt geproduceerd is ook die metadata bekend.

Er is heel duidelijk aan de huidige stand van zaken te zien wat er gebeurt als we dat niet doen!

March 01 2019

From Persistent Identifiers to Digital Objects to Make Data Science More Efficient

Peter Wittenburg

> Author and Article Information

Data Intelligence (2019) 1 (1): 6–21.

https://doi.org/10.1162/dint_a_00004 Article history

quality, transforming formats, etc.). In 2014 M. Brodie reported about a study from Massachusetts Institute of Technology (MIT) [3] according to which 80% of the time of data scientists is wasted on data wrangling and recently (2017) a survey in industry from CrowdFlower [4] reported about 79% of wasted time. In addition, X. Clever reported at the Big Data and AI Summit in 2018^[4] that 60% of industrial data projects fail.

https://doi.org/10.1162/dint_a_00004

[3] = <http://pdfs.semanticscholar.org/46b2/bddce0f3b47f8b9dbe9e05777db9a24d8c39.pdf>

[4] = https://visit.crowdfunder.com/WC-2017-Data-Science-Report_LP.html

healthRI
enabling data driven health

Omdat dit in de praktijk weinig gebeurt moeten projecten die data nu "hergebruiken" deze data eerst leren begrijpen. Dat kost volgens eerder onderzoek tot wel 80% van de tijd van de onderzoekers in zulke projecten.

THEME ARTICLE: SPECIAL SECTION ON THE 75TH ANNIVERSARY OF THE IEEE COMPUTER SOCIETY

75 Years of Astonishing Evolution of IT: 1946-2021

George Strawn¹ US National Academy of Sciences, Engineering, and Medicine

prescription not an implementation. As mentioned in the previous paragraph, currently data is "munged" manually for analysis rather than by automated processes. This results in data preparation/munging taking 80% of the time leaving only 20% for analytics (science or business). I hope/believe that in the decade of the 2020s, we will learn how to automate the process of making data FAIR, which will go a long way to making data Fully AI-Ready much more rapidly.

<https://doi.org/10.1109/MITP.2021.3070963>

healthRI
enabling data driven health

Als we de tijd die aan deze overhead wordt besteed kunnen halveren, komt er 3x zoveel tijd om de echte analyse uit te voeren!

Reusing data takes effort

- Much value from data comes from combining sources
- Data sources are encoded using
 - Non-standards
 - Self-declared *de-facto* standards
 - Different standards

In de praktijk zien we vaak dat er geen standaarden worden gebruikt. In het geval van grote data resources gaat men er vaak van uit dat ze zelf de standaard “zetten”. Ik kan een eenvoudig voorbeeld laten zien dat aangeeft hoe dit in de praktijk opduikt.

Reusing data requires understanding it

Subject-ID	Date of visit	Sex
Patient 1	23 September 2020	Male
Patient 2	22-10-2019	F

Imagine having to deal with this:

- For more than 2 data sources
- For more than 2 columns
- With non-standards

Hier is een voorbeeld van een gecombineerde dataset waar niet is nagedacht over standaarden. We kunnen alle Velden eenvoudig begrijpen, maar ze zijn niet zonder meer programmatisch met elkaar te vergelijken.

Als je dit probleem uitbreidt naar 100 kolommen en 5 databronnen, kun je je voorstellen dat er meer dan een paar minuten in gaat zitten om die data met elkaar koppelbaar te maken. Vooral uitzonderingssituaties kunnen heel veel tijd kosten (denk bijvoorbeeld aan gennamen die door EXCEL in datums worden omgezet). Vandaar die 80% van de tijd die daaraan opgaat.

Reusing data requires understanding it

Subject-ID	Date of visit	Sex
Patient 1	2020-09-23	1
Patient 2	2019-10-22	2

ISO 8601

ISO 5218

Dezelfde data met gebruikmaking van standaarden is veel makkelijker te lezen voor een machine, als tenminste goed is gedocumenteerd welke standaard is gebruikt.

Er zijn iso standaarden voor datums zowel als sekse.

Choose a standard, do not make one up!

- ISO 5128 “does not provide codes for sexes that may be required in specific medical and scientific applications or in applications that need to code sex information other than for human beings”
- Don’t make something up; choose an alternative such as SNOMED-CT.
- Document the choice.
- Mappings between standards are possible and often reusable.

Niet zelf gaan zitten klooiën! Er zijn al heel veel standaarden, er past er vaak een wel!

SNOMED-CT options

- ⊖ Finding related to biological sex
 - ⊖ Female
 - ⊖ Baby female
 - ⊖ Finding of sex of baby
 - ⊖ Indeterminate sex
 - ⊖ Intersex
 - ⊖ Male
 - ⊖ Baby male
 - ⊖ Transsexual
 - ⊖ Female-to-male transsexual
 - ⊖ Female to male transsexual person on hormone therapy
 - ⊖ Surgically transgendered transsexual, female-to-male
 - ⊖ Male-to-female transsexual
 - ⊖ Male to female transsexual person on hormone therapy
 - ⊖ Surgically transgendered transsexual, male-to-female
 - ⊖ Previously asexual transsexual
 - ⊖ Previously bisexual transsexual
 - ⊖ Previously heterosexual transsexual
 - ⊖ Previously homosexual transsexual
- ⊖ Gender finding
 - ⊖ Feminine gender
 - ⊖ Gender unknown
 - ⊖ Gender unspecified
 - ⊖ Masculine gender
 - ⊖ Non-binary gender
 - ⊖ Surgically transgendered transsexual
 - ⊖ Surgically transgendered transsexual, female-to-male
 - ⊖ Surgically transgendered transsexual, male-to-female
 - ⊖ Transgender identity

How?

Eigenlijk zijn we met deze voorbeelden al een beetje van het “Waarom” van FAIR data overgelopen naar het “Hoe”. Nu wil ik daar even een wat breder overzicht van geven, door even de 4 letters van FAIR langs te lopen en te kijken wat er praktisch kan worden gedaan.

Findable

- Choose Repository
 - Domain specific
 - Institutional
 - National
 - Special
- Make sure to get and use Persistent Identifiers
- Describe what the data is, not only why you collected it
- Register the data in a catalogue, where others would look for it

Het stuk over het kiezen van een repository belicht ik vaak in verhalen aan onderzoekers die resultaten FAIR moeten maken. Voor Cohort- en populatiestudies is het duidelijk dat er een zelf-ingericht repository, een eigen web site voor wordt gebruikt. Voor de vindbaarheid is het heel relevant dat er PIDs worden gebruikt. Een PID is een GUPRI: Wereldwijd-unieke naam voor iets die nooit meer voor iets anders zal worden gebruikt en die zelf geen herkenbare metadata bevat, en die een mogelijkheid biedt naar de eigenlijke locatie van een document te gaan (zelfs als die locatie over 15 jaar is veranderd).

Keywords moeten ook worden gekozen op basis van waar onderzoekers naar zullen zoeken, beschrijven WAT data IS is daarvoor effectiever dan de originele **reden** om de data te verzamelen. Beschrijf bijvoorbeeld dat er fasecontrast Röntgenfoto's van het torso in de collectie zitten, in plaats van "foto's van de ruggenwervel". Bibliotheacarissen hebben daar kijk op.

Vindbaarheid vraagt ook dat de collectie kan worden gevonden daar waar onderzoekers zoeken naar dergelijke informatie, in 1 of meer catalogi.

=====

- * Data needs to come into a repository
- * Special = Made for the project. Note Accessibility!
- * Different repositories for different data?
- * Preferences from funder or institute? Budget affected!
- * Go talk to repositories! They may require format, metadata. They can help!

- * Persistent identifier like DOI, Help you get credit

- * Especially institutional or special: catalog it!
- * Keywords for re-use. What is in there, not why you collected
- * Keywords for subsets?
- * Librarians/archivists can help! Go look for them!
- * Talk to the catalog people early to find out what they need.

Accessible

- Assure Longevity
- Check Legal Conditions
- Limit the Embargo
- Use standard protocols also for machine actionability
- Ensure proper ICT procedures

- Levensduur: 20% weblinks vervallen per jaar. De PIDs zijn een deel van de oplossing omdat die veranderingen kunnen volgen. Zorg daarnaast voor een backup-plan: wat gebeurt er met de collectie als je er zelf mee stopt? Zou de data dan ergens anders een plek kunnen vinden? Het "Core Trust Seal" voor repositories vraagt om zulke plannen. FAIR vraagt expliciet om te regelen dat de metadata blijft bestaan zelfs als de data weg is.
- Omdat we eigen repositories maken heb ik hier toegevoegd standaard toegangsprotocollen te gebruiken. Software tools moeten niet speciaal voor elke dataset hoeven te worden hergeprogrammeerd!
- Denk ook aan de levensduur van Software
- Laat professionals zorg dragen voor je backups en beveiliging.

=====

- Longevity: 20% decay per year on links to web sites. Certified repositories! Transfer if they stop.
- Software longevity
- Repositories often work with one-time-fee.
- Privacy sensitive data: Not coupled to person! Clear rules! Find a data access committee
- Non-legal reasons: Limited time! not "once we get the patent" or "after the publication"
- ICT procedures: convince reviewer that you won't lose (track of) the data. Professional?

Interoperable

- Terminology
- Format
- Sufficient metadata

Brussels ← → Bruxelles
Cancer ← → Malignant Neoplasm

Interoperabiliteit bestaat uit 3 onderdelen: niet alleen de terminologie-standaarden die ik als voorbeeld al had genoemd onder “Waarom”, maar ook het formaat waarin de (meta)data beschikbaar is, en de beschrijving van de data, zo precies mogelijk. Probeer te voorkomen dat metadata wordt “weggelaten” terwijl hij er wel is! Vooral voor onverwachte typen hergebruik kan dat toch van belang zijn. Voor het standaardiseren gaat het niet alleen om de “biologische” concepten. Hergebruik van data vereist soms inter-disciplinaire koppelbaarheid.

Hergebruikers moeten geen moeite hebben de data verkeerd te begrijpen!

- 1854 London Cholera outbreak, John Snow made a map and identified a water pump that was the origin of the infection
- Arabidopsis study where correlation was made with height through google maps API.

====

- * Format
- * Terms: (controlled) vocabulary. Or an (international) coding system. With relations: Ontology.
- * Examples after click
- * Use what others use, or map
- * All fields: even if you do genetics, store sample locations like climate scientists.

There will be lots of development, demands on interoperability and benefits will grow!

Example

- 1854 London Cholera outbreak, John Snow made a map and identified a water pump that was the origin of the infection
- Arabidopsis study where correlation was made with height through google maps API.

- You need sufficient metadata to be able to interpret the data

Reusable

- Document provenance
- Use *minimal metadata* standards
- Choose a liberal license
- Make sure results are not only narrated

- Herbruikbaarheid is sterk afhankelijk van metadata.
- Het is essentieel om de data herkomst zorgvuldig te documenteren, dit minimaliseert ook vragen om verduidelijking.
- Als je een minimal metadata standaard kiest, vul die dan maximaal in. Veel van die informatie heb je eigenlijk wel, maar gooi je misschien standaard weg. Zonde!
- Voor de herbruikbaarheid is het belangrijk dat de condities eenvoudig zijn. Probeer zo open mogelijk te zijn en zoveel mogelijk alleen wettelijke beperkingen op te leggen; anders maak je het hergebruikers heel lastig bij het combineren van databronnen.

====

- Annotate to help self and others, avoid clarification requests
- How was data obtained, provenance. Instrument settings. Automate collection! Helps yourself!
- Minimal metadata. Quite extensive, and volunteer to do optional too. Field specific+DC.
- License. Are you obliging people to Cite? Forbid commercial use? Such clauses have Consequences. Unnecessary restrictions!
- See the example of the Gauss curve. Which nobody cites any more!
- Ewan Birney insight: other scientists will cite, they need to prove they got data from a reliable source!

Do not encode knowledge only in narrative. Please help to make text mining obsolete.

* use experts, be experts.

FAIR Data Maturity Model

- For the EC, the Research Data Alliance built a “FAIR Data Maturity Model”
 - <https://doi.org/10.15497/RDA00050>
- Series of yes/no checks you can do on your data
- Some tools exist to do this automatically (subsets of the full model)
- It is not important to measure how FAIR compliant you are, but to identify how you could easily become more FAIR in the future

Een model is prima om heel duidelijk te maken wat de FAIR principes in de praktijk betekenen. Maar het is zinloos om blind een score te laten uitrekenen, want die scores zijn onvergelijkbaar. Het is veel beter om te kijken hoe je je data eenvoudig kunt veranderen met het grootste effect op de FAIRness. Zie het als investeringsanalyse.

A model like this is very good to make very clear what the original principles mean. But it is useless to do automatic and blind checking to get a “score”, because they are incomparable! Instead, it is better to check where you can make changes to your data that are easy to do and valuable for FAIRness. An investment analysis.

Starting the FAIR journey is doable!

Er is door Cameron Neylon al snel gezegd: “FAIR is een reis, geen reisdoel”. Ik heb hier suggesties gegeven voor de eerste stappen op die reis. En ik geef de suggestie om je ook in de toekomst steeds te blijven afvragen of het eenvoudig nog FAIRder kan.

The presentation will start soon



This webinar will be recorded

Recordings can be made available for knowledge dissemination.



Questions or remarks?

Please raise your hand by clicking on the “raise your hand”-symbol or use the chat.



Mute your microphone

Please mute your microphone to minimize background noise by clicking on the “mute”-symbol on the top of your screen.



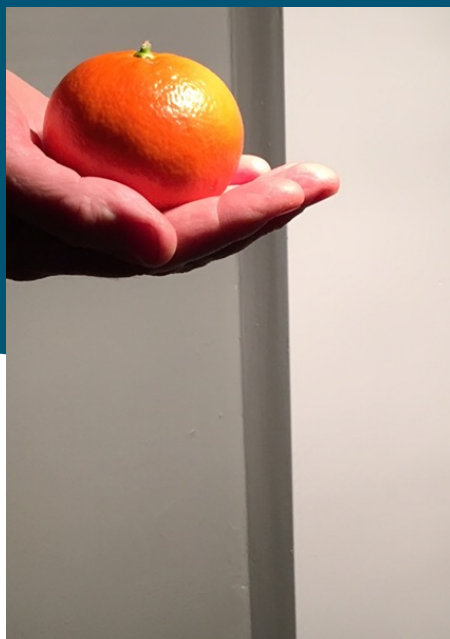
What can we improve for the future?

We are happy to hear any additional questions or feedback for our next expert meeting. Please contact us by mail: robin.verjans@lygature.org

BBMRI-workshop #1, 9 september 2021

Data genereren volgens de FAIR principes

About FAIRifying your data in ZonMw-funded projects



Margreet Bloemers, ZonMw

Projectleider FAIR data en datamanagement

toegangtotdata@zonmw.nl



What is ZonMw?

The Netherlands Organisation for Health Research and Development

<https://www.zonmw.nl/en/>

- ZonMw stimulates health research and health care innovation
- ZonMw funds health research and stimulates the application of the new knowledge – to improve health and (health) care.

Impact from research output

News and funding

Research and results

- › Efficiency Studies
- › Fundamental research
- › Handicap and Chronic Disease
- › Infectious Diseases and Antimicrobial Resistance
- › Life Sciences & Health
- › Mental Health
- › Medicines
- › Palliative care
- › Prevention
- › Quality in Health Care
- › Sport
- › The Elderly
- › Translational research
- › Youth

Policy Points

- › Co-financing
- › FAIR data & data management
- › Positive health
- › Societal impact (Implementation)

Outline & learning outcomes

1. Why does a research funder bother about FAIR data and research data management RDM?



2. What is the funder's role? How can the funder play this role to achieve its goals?



3. Corona and the leap forward in FAIRification:
from F & A to I & R



Ad1 - Why RDM & FAIR data?

Preconditions for data reuse and data sharing!

□ Enable data reuse and data sharing:

- Allow other researchers to build upon existing data and knowledge (return on investment)
- Accelerate knowledge and innovation for improving health
- Enrich research (combining datasets)
- Pattern recognition in large datasets
- Return on investment



Data and other research outputs, e.g.:

Biological materials, recordings, images, etc.

□ Enable verification of research results:

- Research integrity, responsible science
- Reproduction research
- Quality

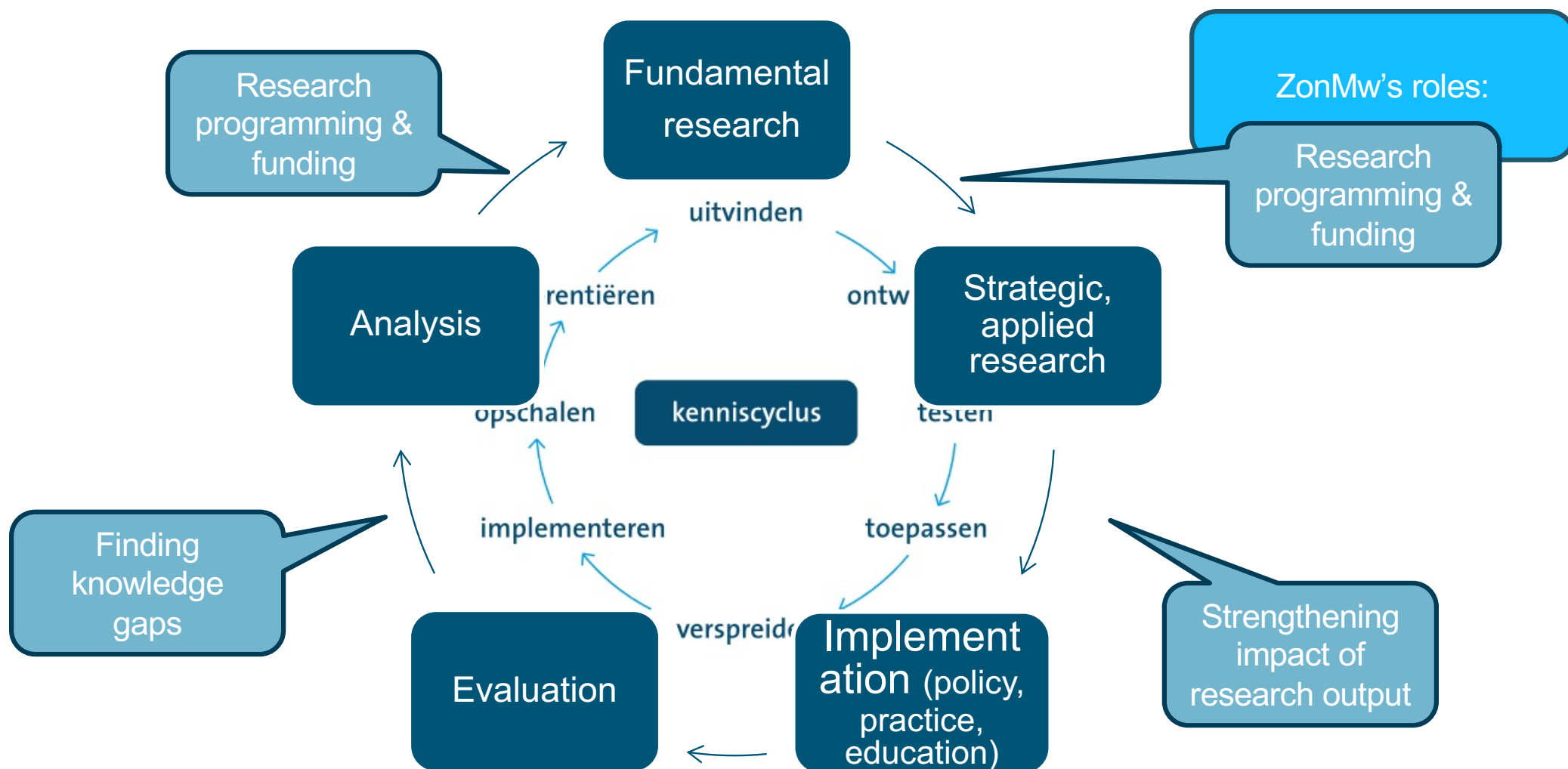


□ Ensure added value for the researcher!

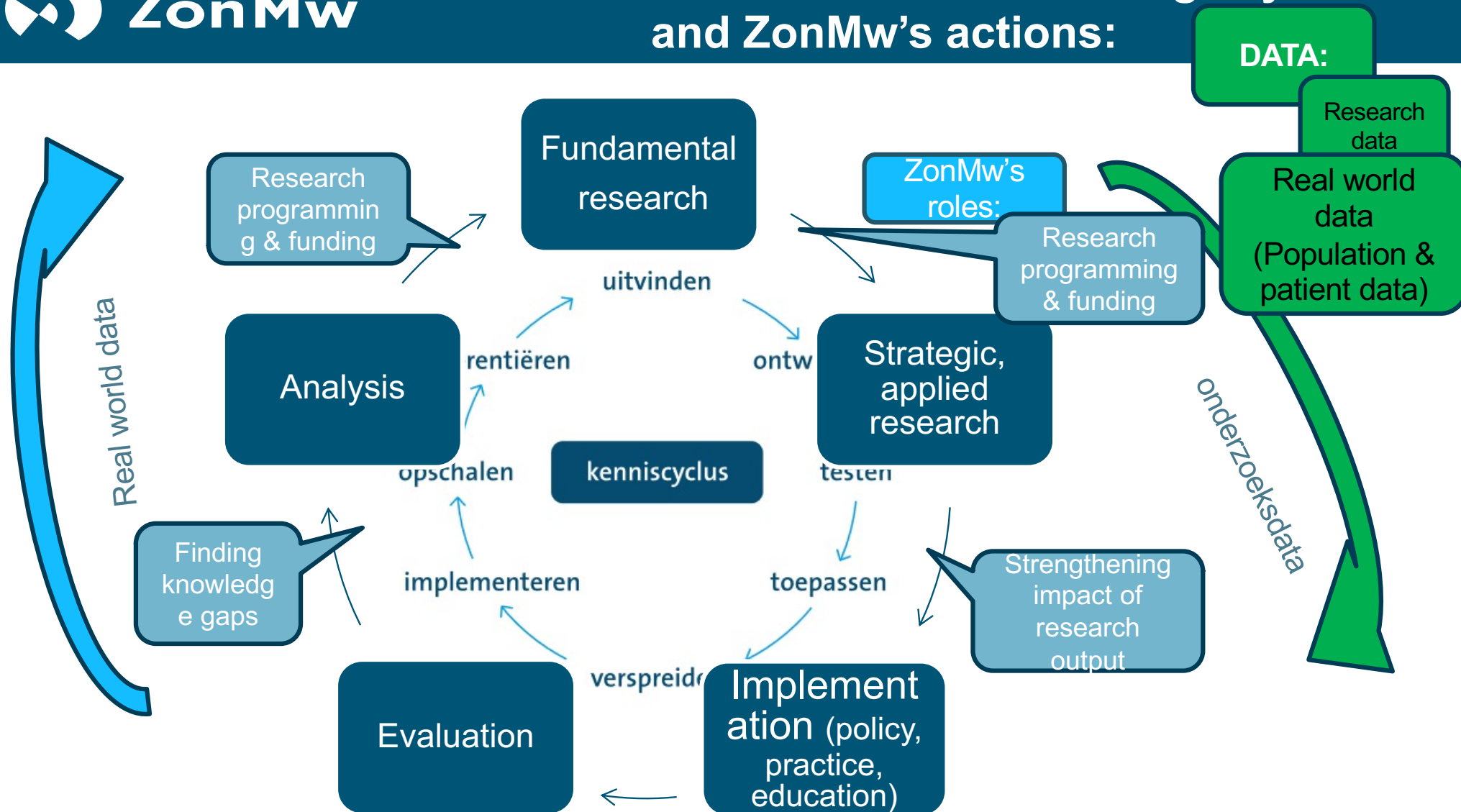
1. Data management (stewardship)

2. Data/assets FAIR maken!

The relationship between the knowledge cycle and ZonMw's approach towards innovation:



Data inform the knowledge cycle and ZonMw's actions:



Outline & learning outcomes

1. Why does a research funder bother about FAIR data and research data management RDM?
2. What is the funder's role? How can the funder play this role to achieve its goals?
3. Corona and the leap forward in FAIRification: from F & A to I & R



Ad 2 – What is the role of the funder in achieving FAIR/reusable data? ZonMw's examples:



☐ As a funder:

Setting requirements for grant proposals and awarded projects to deliver 'FAIR' data

☐ As a facilitator:

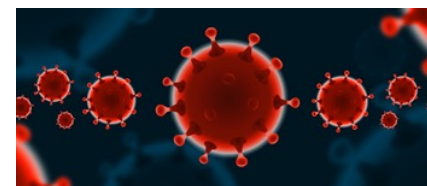
Enabling development of community specific FAIR data practices
(e.g. FAIRification workflow)

☐ As a consortium partner:

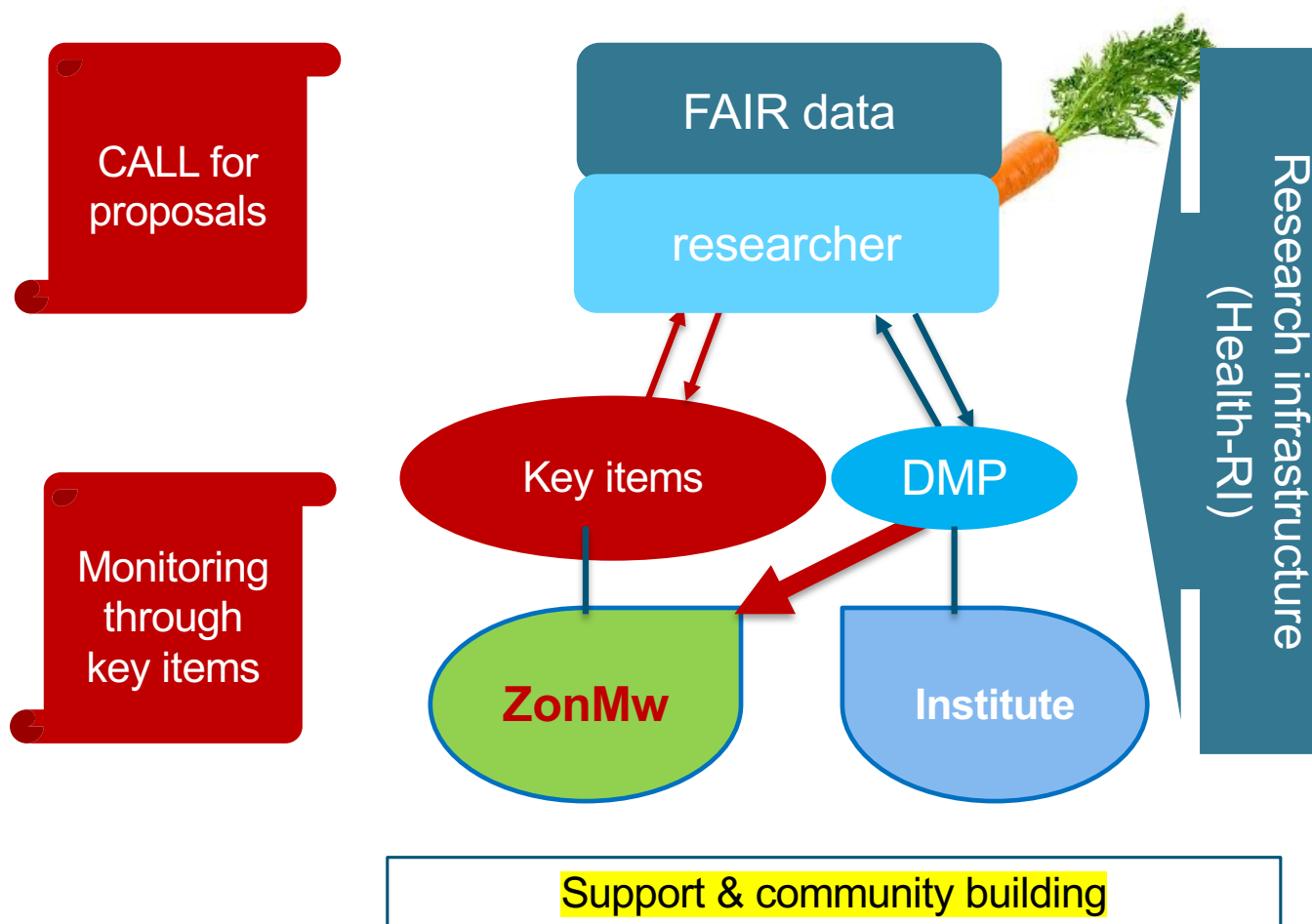
In research & funding networks: putting the insights from the FAIRification workflow into practice



The funder as
a driving force!



Ad 2 - ZonMw's approach to data management



Requirements & Support:

- RDM – DMP
- Key items about your output
- Involve a datasteward
- Budget for RDM
- Adhere to standards in your community
- Consider M4M (3-PFF)
- Data portal at Health-RI
- Workshops

FAIR principles are a guide towards reusable data

FAIR guiding principles:

Findable	– vindbaar
Accessible	– toegankelijk
Interoperable	– uitwisselbaar
Re-usable	– herbruikbaar

*FAIR data can be **found, understood,**
and reused by humans and by machines
(computers)*

FAIR >> F-AI-Ready

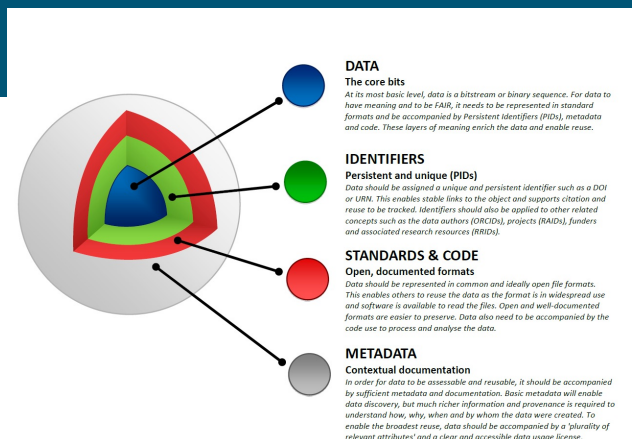


Figure 1: A model for FAIR Data Objects, noting the elements that need to be in place for data to be Findable, Accessible, Interoperable and Reusable.



FAIR is not all or nothing >> FAIRification

WHAT IS FAIR ?

Findable:

F1 (meta)data are assigned a globally **unique and persistent** identifier; **FM-F1A FM-F1B**
F2 data are described with **rich metadata**; **FM-F2**
F3 metadata clearly and explicitly include the identifier of the data it describes; **FM-F3**
F4 (meta)data are registered or **indexed** in a searchable resource; **FM-F4**

Interoperable:

I1 (meta)data use a formal, accessible, shared, and broadly applicable **language for knowledge representation**; **FM-I1**
I2 (meta)data use **vocabularies that follow FAIR principles**; **FM-I2**
I3 (meta)data include **qualified references** to other (meta)data; **FM-I3**

Accessible:

A1 (meta)data are retrievable by their identifier using a standardized communications protocol; **FM-A1.1**
A1.1 the protocol is **open, free, and universally implementable**; **FM-A1.2**
A1.2 the protocol allows for an **authentication and authorization** procedure, where necessary; **FM-A2**
A2 metadata are accessible, **even when the data are no longer available**; **FM-A2**

Reusable:

R1 (meta)data are richly described with a plurality of accurate and relevant attributes; **FM-R1.1**
R1.1 (meta)data are released with a clear and accessible **data usage license**; **FM-R1.1**
R1.2 (meta)data are associated with **detailed provenance**; **FM-R1.2**
R1.3 (meta)data meet domain-relevant **community standards**; **FM-R1.3**

Sci. Data 3:160018 doi: 10.1038/sdata.2016.18 (2016)
<http://fairmetrics.org>
<https://github.com/FAIRMetrics/Metrics/blob/master/ALL.pdf>



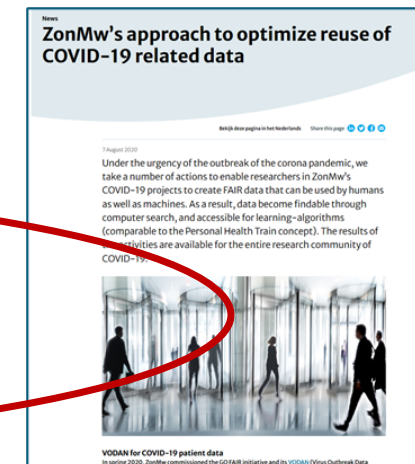
Some important features:

- FAIR is for 90% about metadata
- Machine readable > the computer can find, understand, use
- Domain specific, relevant for community

ZonMw promotes
community agreement!

Outline & learning outcomes

1. Why does a research funder bother about FAIR data and research data management RDM?
2. What is the funder's role? How can the funder play this role to achieve its goals?
3. Corona and the leap forward in FAIRification:
from F & A to I & R



Next steps in FAIRifying data in ZonMw-projects

- ❑ The urgency of Corona

Virus Outbreak Data Network (VODAN)

WHO eCRF data model (semantic, contr.vocabs, machine readable)

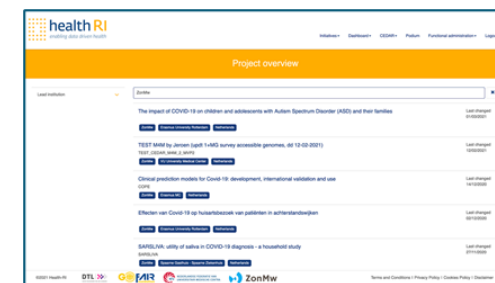
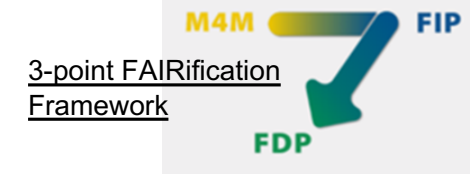
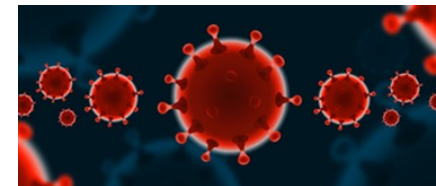
- ❑ Followed by FAIR data services in the COVID-programme (GO FAIR Foundation, DTL, Health-RI)

Some important features:

- FAIR is for 90% about metadata
- Machine readable > the computer can find, understand, use
- Domain specific, relevant for community

ZonMw promotes community agreement!

- Metadata for machines (M4M)
- specific for COVID-research (all subdisciplines)
- Health data portal
- Support & community building

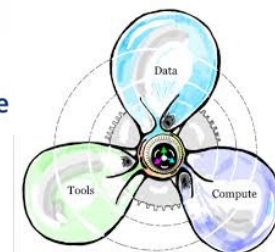


National Find&Access portal

Who do you need?

Datastewards take a central position!

- ❑ They help PIs to FAIRify their data
- ❑ ZonMw requires grantees to involve data stewards in the projects
- ❑ Data stewards obtain specialised knowledge at data services & research infrastructures



Experts at FAIR data services & research infrastructures



Open Data Infrastructure for Social Science and Economic Innovations

Funding agencies

(ZonMw, NWO, Horizon Europe, ...)



researcher



Data Steward

Data steward at univ/umc/hs

Local Digital Competence Centres



M4M.7
13 projects
21 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

M4M.8
10 projects
17 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

M4M.9
6 projects
12 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

M4M.10
7 projects
13 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

M4M.11
9 projects
16 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

M4M.12
14 projects
21 participants

1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

Rob's lecture:

"Choose a standard, do not make one up!"

*"Describe **what the data is**, not why you collected it"*

COVID-19 Program
controlled list
1. Data
2. Images
3. Biomaterials
4. Services
5. Standards

Build & register
ZonMw COVID-19 Program Vocab

Project Admin
Metadata

Project Content
Metadata

Generic Dataset
Metadata

M4M Support Team



FAIR Data Points

Researchers create
Metadata instances
(Routine)

COVID-19 Program
Metadata input forms

**Now also M4M for
ID&AMR:**

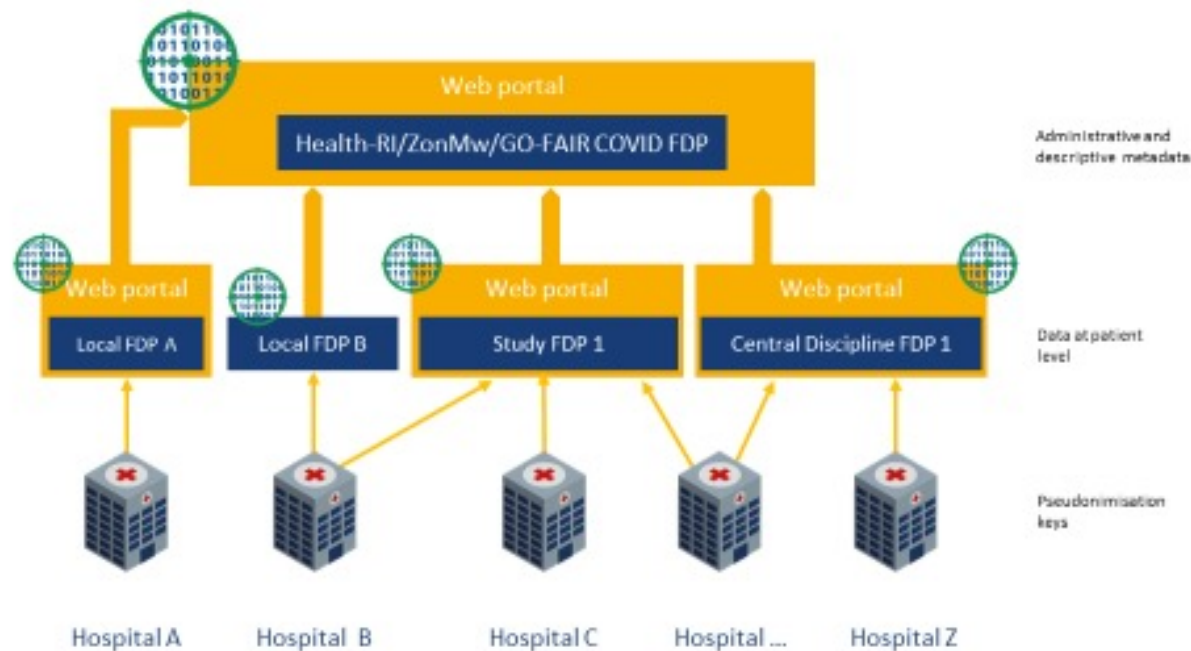
- projects,
- collections of biological materials (BBMRI)

- Complementary to the FAIR Maturity Matrix

>> see Rob's lecture!!



Towards a national COVID-19 observational data portal



FDP- FAIR Data Point

More information about the Health Data Portal at Health-RI:

❑ Workshops about M4M and data portal:

<https://www.health-ri.nl/events/re-play-health-ri-conference-2021>

❑ Slide deck:
COVID data portal and data FAIRification
2021-03-04 Health-RI conference

❑ Kick off meeting Health-RI
Groeifonds October 5, 2021



WARNING DISCLAIMER URGENCY!!!

- ☐ We need MORE datastewards!
- ☐ Who are eager to learn about 3-PFF
- ☐ And become M4M champions
- ☐ Don't forget: the one who generates the data, is the first datastewards!

<https://docs.google.com/presentation/d/1Gdt99TKW-IMRWaL73YLH15IXEqu80Rx7P1PfbzCuiDw/edit#slide=id.p2>



Want to know more?

ZonMw

FAIR data and data management:

<https://www.zonmw.nl/en/research-and-results/fair-data-and-data-management/>

Margreet Bloemers

Lisette Schuur

Bas de Waard

ToegangTotData@zonmw.nl

OpenScience@zonmw.nl

OPEN DISCUSSION

1. Please take your mobile device
2. Go to www.menti.com
3. Use the code **24 74 90 4**
4. Participate and share your thoughts!

<https://www.menti.com/nr859os3tp>

OPEN DISCUSSION

- Please take your mobile device
- Go to *www.menti.com*
- Use the code **24 74 90 4**
- Participate and share your thoughts!

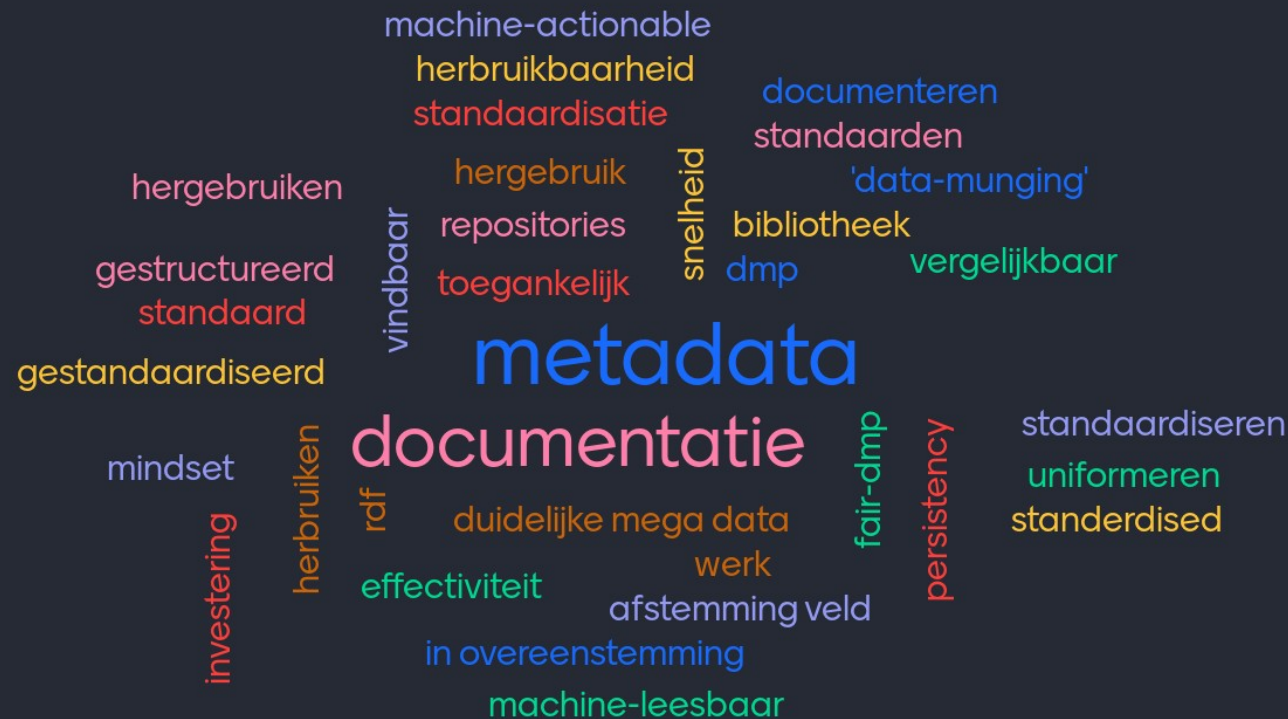


Doel van de discussie:

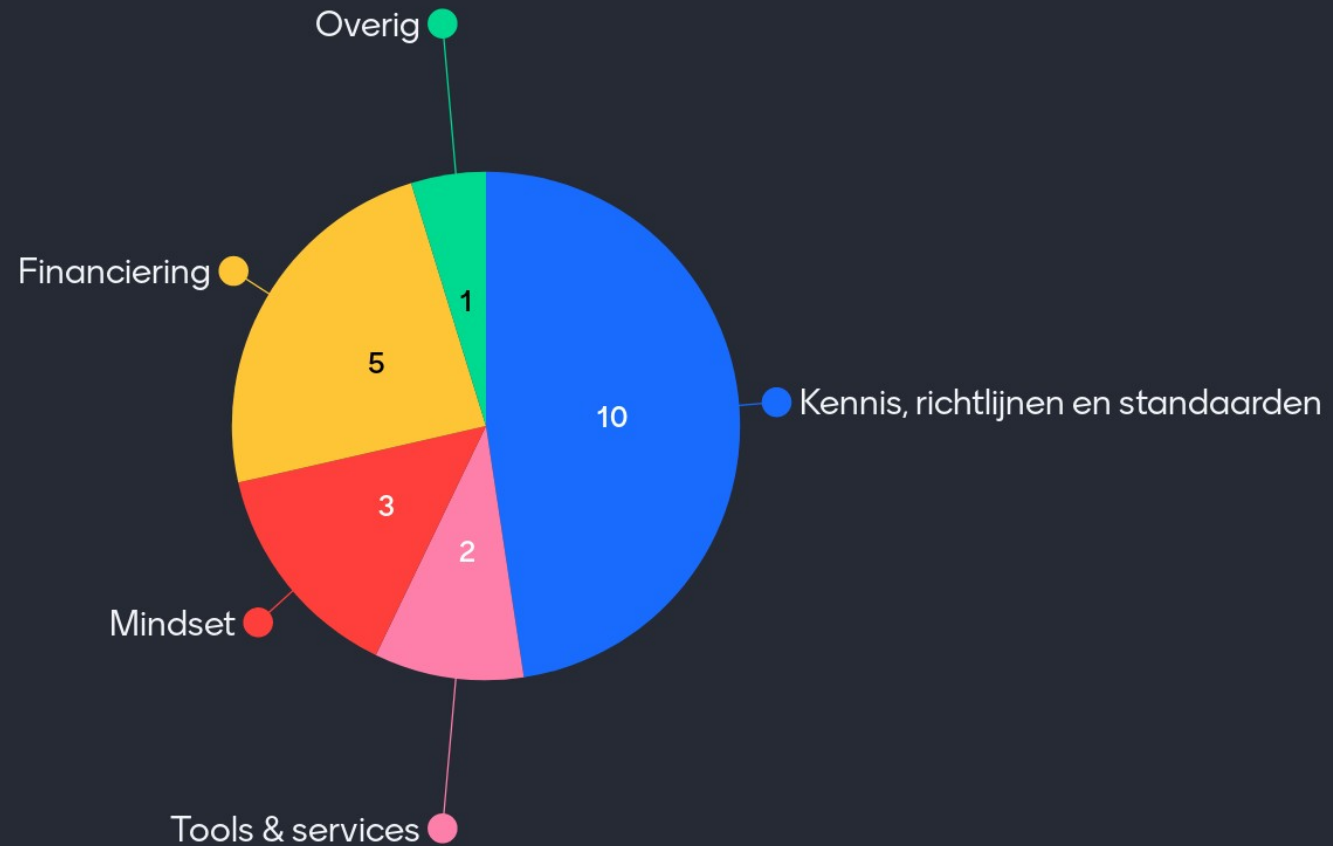
- 1) Wat moet er gebeuren om over 5 jaar data FAIR op te leveren? (Toekomstvisie)
- 2) Welke stappen moeten hiertoe als eerste genomen worden om tot deze toekomstvisie te komen? (Low-hanging-fruit)



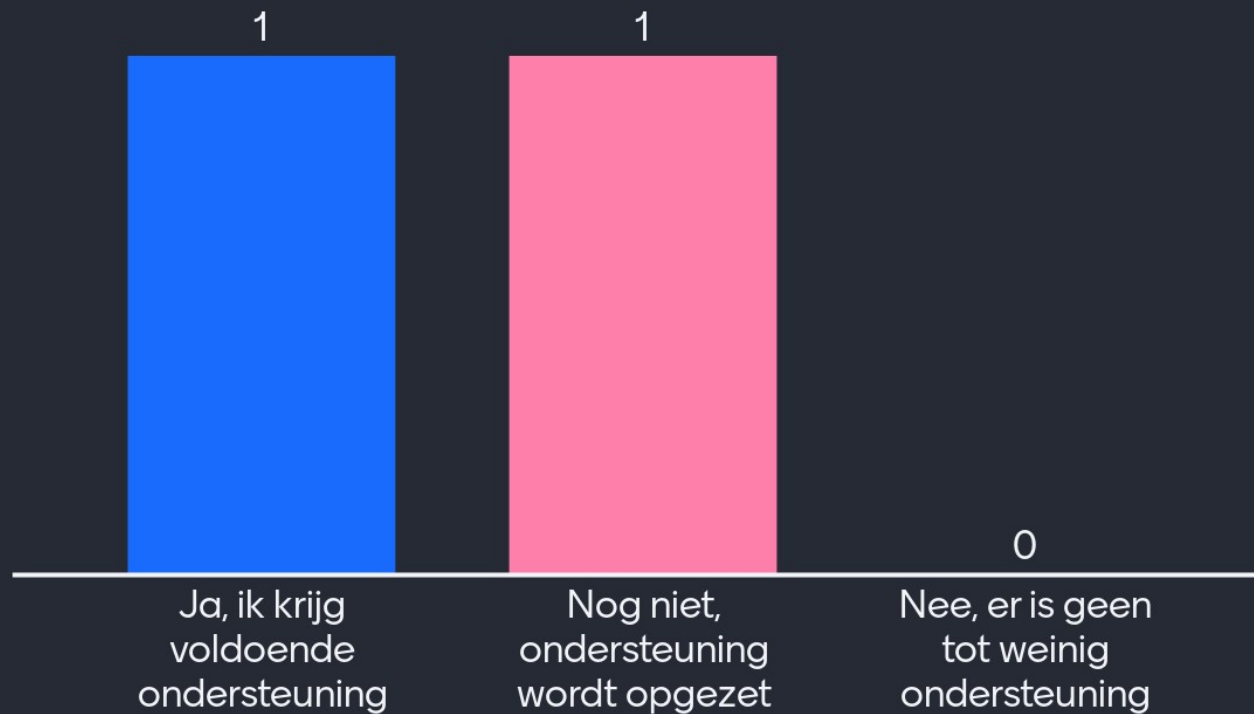
1. Wat houdt het in om data FAIR te genereren? (één-woord antwoorden)



2. Welke ondersteuning is noodzakelijk om data FAIR te genereren?



3. Krijgen jullie voldoende ondersteuning vanuit jullie instelling/werkgever voor het FAIR-genereren van data?



4. Welke rol moet de subsidieverstrekker spelen bij het beschikbaar maken van data?



- Geen, de data-aanbieder moet de vrijheid hebben
- Het geven van onderwijs in de FAIR-principes
- FAIRificatie van data als harde eis voor (gedeeltelijke) subsidieverstrekking
- Overig



5. Welke stappen moeten eerst gezet worden om van de huidige situatie tot het FAIR genereren van data te komen?

DMP

Bruikbare metadata templates

Er zijn al bestaande initiatieven (o.a. van softwareontwikkelaars) om data FAIR te maken 'bij de bron'. Opnieuw een infrastructuur opzetten per instituut zou zonde zijn en alleen maar meer tijd / andere resources kosten.

Kennis, richtlijnen en services delen en beschikbaar maken

zo breed mogelijk gedragen standaarden, liefst internationaal

Kijken naar best practices in Nederland en die actief promoten.

Lokaal goede informatievoorzieningen, zoveel mogelijk verwijzend naar landelijke/globale standaarden en oplossingen

Early career researchers maar zeker ook studenten zijn key. Senior onderzoekers meenemen is goed maar als ze hun studenten laten freewheelen mis je veel fair potentie.

Middelen (mankracht, financiën en tools) om bestaande data om te zetten naar FAIR standaarden

5. Welke stappen moeten eerst gezet worden om van de huidige situatie tot het FAIR genereren van data te komen?

Duidelijke voorbeelden bieden, voor velen is FAIR nu te abstract en klinkt het vooral als veel werk.

Voorbeelden

WE HOPE TO SEE YOU NEXT EXPERT MEETING!



Data genereren volgens de FAIR principes

Donderdag 9 september 13.00-14.30 uur

- Dr. Rob Hooft – DTL
- Dr. Margreet Bloemers – ZonMw
- Moderator: Prof. Dr. Wiro Niessen/ Dr. Jan-Willem Boiten



Data vindbaar en koppelbaar maken

Donderdag 23 september 13.00-14.30 uur

- Prof. Dr. Sabine Siesling - IKNL/PSCCR register
- Prof. Dr. Edwin Cuppen - HMF/GENONCO
- Moderator: Dr. Jeroen Beliën - Amsterdam UMC



Data zoeken en toegankelijk maken

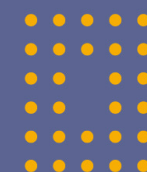
Donderdag 7 oktober 13.00-14.30 uur

- Dr. Aletta Debernardi - PALGA
- Prof. Dr. Edwin Cuppen – HMF
- Moderator: Prof. dr. Folkert van Kemenade – Erasmus MC

HEALTH-RI MEETING

GEARED UP TO ACCELERATE

05 OCT 2021
JAARBEURS MEDIA PLAZA
UTRECHT



health RI
enabling data driven health

Join us on **Tuesday 5 October** for the **Health-RI Accelerator meeting!**
Jaarbeurs Media Plaza in Utrecht

Festive kick-off with all parties involved. All ready and geared up to
realize a national health data-infrastructure for research and
innovation.

Looking forward to meeting you all 'in real life' again!

PROGRAM INFO

MEDIA PLAZA INFO

