



Ministerie van Binnenlandse Zaken en
Koninkrijksrelaties

principes FAIR data

toegepast in de Zorg en bij BZK en VRO

met focus op interoperabiliteit van (open) standaarden

Barend Mons (GO FAIR Foundation, LIFES), Yola Park (Nictiz), Rens Kievit (BZK, LUMC),
Norbert van Dijk (BZK, VRO)

Dag van de Interoperabiliteit





Agenda

1. Introductie FAIR, GO FAIR Fellowship, eerste stappen met FAIR bij BZK en VRO, Norbert van Dijk, CIO-adviseur, BZK en VRO
2. Promotieonderzoek semantische toegangscontrole en FAIR, Rens Kievit, PhD candidate BZK en LUMC
3. FAIR en databeschikbaarheid in de zorg, Yola Park, teamlead Stelselbeheer, Nictiz
4. Semantic types en open standaarden, Barend Mons, emeritus hoogleraar Biosemantiek, LUMC en Universiteit Leiden, oprichter en wetenschappelijk directeur GO FAIR Foundation, oprichter Leiden Institute for FAIR and Equitable Science LIFES



FAIR

- > oorsprong wetenschappelijk onderzoek
- > toepassing principes maken data vindbaar, toegankelijk, interoperabel en herbruikbaar
- > toepassing principes maken data ook "machine readable", "machine actionable", "Fully AI Ready"
- > essentiele rol voor metadata

SCIENTIFIC DATA

Amended: Addendum

OPEN **Comment: The FAIR Guiding Principles for scientific data management and stewardship**

Mark D. Wilkinson *et al.*[#]

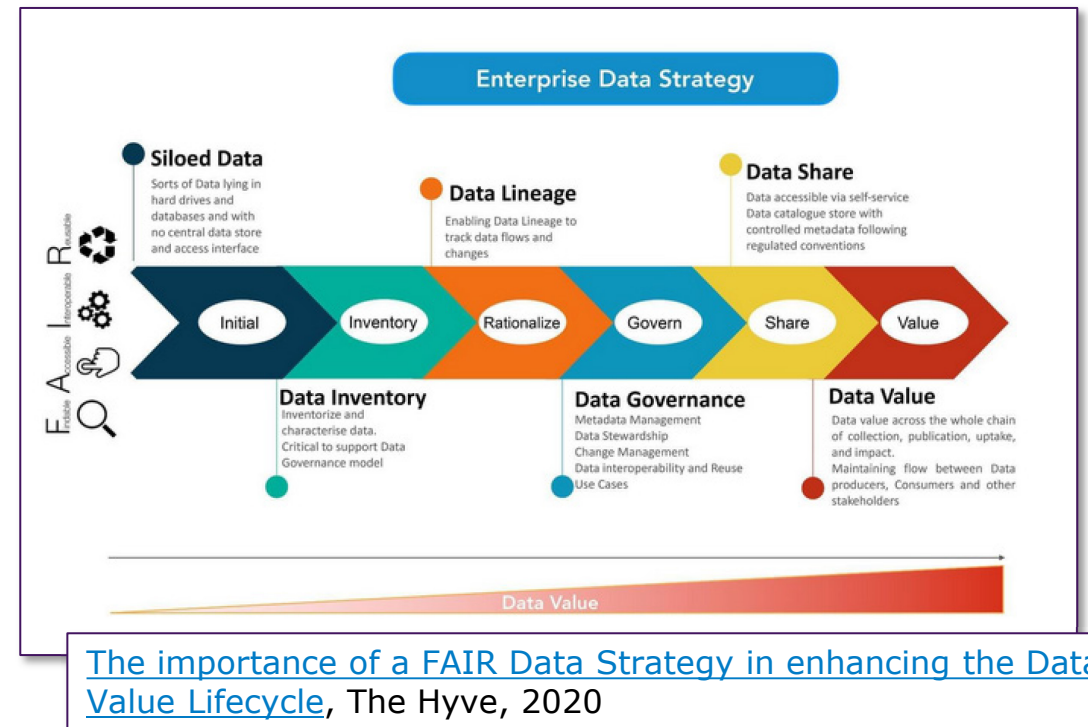
Received: 10 December 2015
Accepted: 12 February 2016
Published: 15 March 2016

There is an urgent need to improve the infrastructure supporting the reuse of scholarly data. A diverse set of stakeholders—representing academia, industry, funding agencies, and scholarly publishers—have come together to design and jointly endorse a concise and measureable set of principles that we refer to as the FAIR Data Principles. The intent is that these may act as a **guideline for those wishing to enhance the reusability of their data** holdings. Distinct from peer initiatives that focus on the human scholar, the FAIR Principles put specific **emphasis on enhancing the ability of machines to automatically find and use the data**, in addition to supporting its reuse by individuals. This Comment is the first formal publication of the FAIR Principles, and includes the rationale behind them, and some exemplar implementations in the community.



Toegevoegde waarde

- > FAIR data
 - vindbaar
 - toegankelijk
 - interoperabel
 - herbruikbaar
- > in een in toenemende mate gedistribueerd data landschap
 - data bij de bron
- > FAIR data is "machine actionable"
 - inzetbaar voor AI initiatieven
- > Nota bene
 - FAIR data $\neq$ open data





Principes

Dubbeltklik om ingevoegd document te openen.

FAIR Principles



In 2016, the 'FAIR Guiding Principles for scientific data management and stewardship'¹ were published in Scientific Data. The authors intended to provide guidelines to improve the Findability, Accessibility, Interoperability, and Reusability of digital assets. The principles emphasise machine-actionability (i.e., the capacity of computational systems to find, access, interoperate, and reuse data with none or minimal human intervention) because humans increasingly rely on computational support to deal with data as a result of the increase in volume, complexity, and creation speed of data.

Findable

The first step in (re)using data is to find them. Metadata and data should be easy to find for both humans and computers. Machine-readable metadata are essential for automatic discovery of datasets and services, so this is an essential component of the FAIRification process.

- F1. (Meta)data are assigned a globally unique and persistent identifier.
- F2. Data are described with rich metadata (defined by R1 below).
- F3. Metadata clearly and explicitly include the identifier of the data they describe.
- F4. (Meta)data are registered or indexed in a searchable resource.

Accessible

Once the user finds the required data, she/he/they need to know how they can be accessed, possibly including authentication and authorisation.

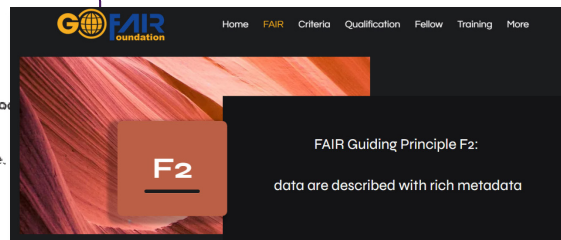
- A1. (Meta)data are retrievable by their identifier using a standardised communications protocol.

A1.1 The protocol is open, free, and universally implementable.

A1.2 The protocol allows for an authentication and authorisation procedure where necessary.

- A2. Metadata are accessible, even when the data are no longer available.

¹ See <https://www.nature.com/articles/sdata201618>





Interpretations



Interpretations

The FAIR Guiding Principles [1] provide guidance when improving Findability, Accessibility, Interoperability and Reusability of digital research data. The FAIR Foundation believes that, where possible, interoperability should be achieved through the use of standards that are widespread and consistent. The community of FAIR researchers and practitioners aim is to provide a reference for the implementation of machine-actionability. The GO FAIR Foundation

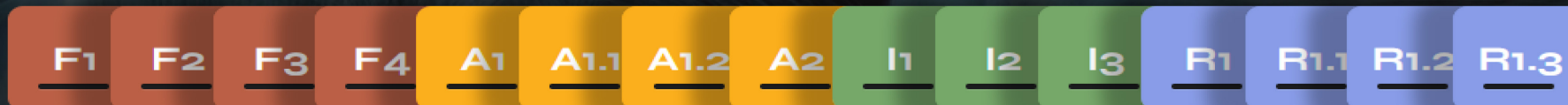
Volume 2, Issue 1-2
Winter-Spring 2020



January 01 2020

FAIR Principles: Interpretations and Implementation Considerations

Annika Jacobsen, Ricardo de Miranda Azevedo, Nick Juty, Dominique Batista, Simon Coles, Ronald Cornet, Mélanie Courtot, Mercè Crosas, Michel Dumontier, Chris T. Evelo, Carole Goble, Giancarlo Guizzardi, Karsten Kryger Hansen, Ali Hasnain, Kristina Hettne, Jaap Heringa, Rob W.W. Hooft, Melanie Imming, Keith G. Jeffery, Rajaram Kaliyaperumal, Martijn G. Kersloot, Christine R. Kirkpatrick, Tobias Kuhn, Ignasi Labastida, Barbara Magagna, Peter McQuilton, Natalie Meyers, Annalisa Montesanti, Mirjam van Reisen, Philippe Rocca-Serra, Robert Pergl, Susanna-Assunta Sansone, Luiz Olavo Bonino da Silva Santos, Juliane Schneider, George Strawn, Mark Thompson, Andra Waagmeester, Tobias Weigel, Mark D. Wilkinson, Egon L. Willighagen, Peter Wittenburg, Marco Roos, Barend Mons , Erik Schultes





Implementaties

- > FAIR Enabling Resource Types
 - Architecture Building Blocks

FAIR Sub-principle	Type of FAIR Enabling Resource	Definition
F1	identifier service	A service that provides for any digital object (1) algorithms guaranteeing global uniqueness, (2) policy document that guarantees persistent and (3) resolution of the identifier to machine-actionable metadata describing the object and its location.
F2	metadata schema	A specification that defines metadata fields describing attributes of data or other digital objects.
F3	metadata-Data linking schema	A specification that provides a unique, persistent, (ideally) bi-directional, machine-actionable link between metadata and the data they describe.
F4	registry	A service that indexes metadata and data and provides search over that index.
A1.1	communication protocol	A specification that defines how messages are structured and exchanged.
A1.2	authorization and authentication service	A service that mediates access to digital objects according to specified conditions.
A2	metadata preservation policy	A document that describes the conditions under which metadata are to be provisioned in the future (maybe part of a data management plan).
I1	knowledge representation language	A language specification whereby knowledge can be made processible by machines.
I2	structured vocabulary	A specification of uniquely identified and unambiguous concepts with their definitions represented preferably using web standards.
I3	semantic model	A specification that defines qualified relations between entities describing data or other digital objects using structured vocabularies.
R1.1	usage license	A document that describes the conditions under which a digital object can be legally used.
R1.2	provenance model	A specification that defines metadata fields describing the origin and lineage of data or other digital objects.
R1.3	FIP as a whole	

I2	structured vocabulary	A specification of uniquely identified and unambiguous concepts with their definitions represented preferably using web standards.
----	-----------------------	---



Implementaties

- FAIR Enabling Resources
 - Solution Building Blocks

SURF Services Topics

Persistent Identifiers

Your data always findable

Persistent Identifiers (PIDs) ensure that your objects, such as research data, can always be found and remain so. PIDs can be compared to an ISBN number for books. Even if the location or underlying infrastructure changes, the reference remains intact.

FAIR Connect

126 results for **Structured vocabulary**

[Detailed view](#)

ISO 3166-1:2020
October 2, 2024
The International Standard for country codes and codes for their subdivisions. The purpose of ISO 3166 is to define internationally recognized codes of letters and/or numbers that we can use when...

Types: Structured vocabulary

[Detailed view](#)

ISO 3166-1:2013 Countries codes and their subdivisions
October 2, 2024
ISO 3166-1:2013 is intended for use in any application requiring the expression of current country names in coded form; it also includes basic guidelines for its implementation and maintenance. It is...

Types: Structured vocabulary

[Detailed view](#)



Implementatiekeuzes

- FAIR Implementation Profile
 - Technologiekeuzes en andere afspraken binnen een bepaald beleidsveld of onderzoeksdomein

PandemicPACT FIP [Detailed view](#)

Used by: GloPID-R | Global Research Collaboration for Infectious Disease Preparedness

FAIR Connect

5 results for **Social Policy**

AnaEE CREA FIP 2019 [Detailed view](#)

Used by: AnaEE CREA community

AnaEE CREA FIP 2021 [Detailed view](#)

Used by: AnaEE CREA community

[Detailed view](#)

[Up arrow](#)



GO FAIR Foundation

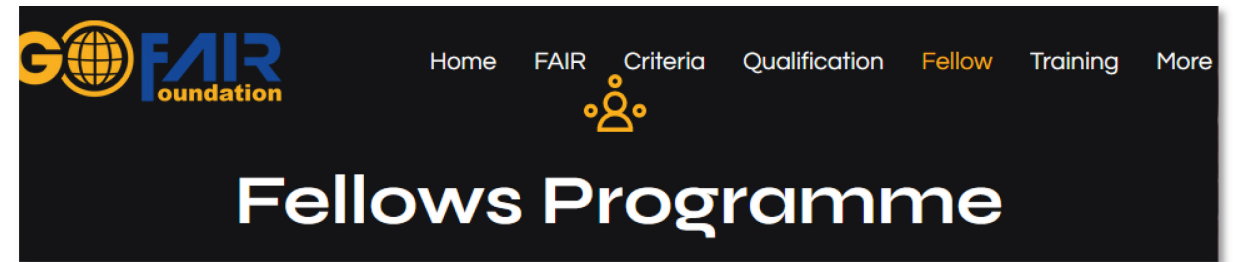
- > hoeder van de FAIR principes en hun interpretaties
- > best practice voor implementatie: 3-Point FAIRification Framework (3PFF)
- > 3PFF Training programme
- > individuele fellowships, maar ook op contractbasis voor (wetenschappelijke) organisaties, bijvoorbeeld National Institutes of Health, US





Ervaringen GO FAIR Fellow

- > inspirerende community, gedreven en gemotiveerde mensen, experts FAIR data management, semantiek en open science
- > focus training ligt op datamanagement in de context van wetenschappelijk onderzoek, research datamanagement
- > schaalniveau onderzoeksproject of programma, wetenschappelijke community, internationaal
- > (farmaceutische) industrie ontwikkelt inmiddels maturity frameworks voor implementatie op organisatieniveau: Pistoia Alliance, Roche
- > deze maturity frameworks lijken in belangrijke mate herbruikbaar voor overheidsorganisaties





BZK & VRO

> Use cases

- "Stedelijke hitte", onderzoeksgegevens delen in de context van klimaatadaptatie van de gebouwde omgeving, gegevenswoorden boek hitte (initiatiefnemer I&W)(VRO)
- Use case open standaarden, toepasbaarheid van semantische types en achterliggende vocabularies op metadata in context overheid

- Onderzoek

- Promovendus, [Rijks I-Doctoraatsprogramma](#), Rijksorganisatie Ontwikkeling, Digitalisering en Innovatie (BZK)

> Beleid

- [Toolbox verantwoord data gebruik](#) (BZK)(Extern)
- [Algoritmekader](#) (BZK) (Extern)
- Gegevensbeleid en AI-beleid (BZK en VRO)(Intern)

> Infrastructuur

- Interne DCAT compliant data catalogus (BZK en VRO)
- Infrastructuur voor duurzaam toegankelijke opslag van data (BZK en VRO)





Maturity frameworks op organisatieniveau

- > Maturity frameworks kunnen de implementatie van FAIR principes op organisatieniveau ondersteunen



Level	Nickname	Key Features
0	Life is unFAIR	Lack of awareness, possibly acquiring awareness
1	Not FAIR	Not FAIR
2	Getting FAIR	Pilots for implementation are in place
3	Pretty FAIR	Transition to good and best practice
4	Really FAIR	Operational, best practice known at the time of writing. Internal organisational focus. Emerging cross-organisation
5	FAIRest of them all	Aspirational. Conceivable but still needs to be realised. Cross-organisation standards and Interoperability

Table FAIR maturity matrix: Maturity Levels.

FAIR Maturity Matrix V1.0 © 2024 by [Pistoia Alliance](#) is licensed under. [CC BY 4.0](#)



Level 1 - "Started the FAIR journey"



- > **data** .. some data is cataloged.. no domain model is used .. heterogeneity is a feature of this stage .. emerging awareness and application of metadata
- > **leadership** ..leadership awareness of FAIR data starts .. awareness of FAIR data starts, and there is some understanding of the (potential) value of FAIR data .. combination of "top-down" or "bottom-up" leadership .. C-suite leaders may lack the specific knowledge .. may have some buy-in
- > **strategy** .. awareness of FAIR data started .. vision and plan for FAIR data begins to emerge .. shift towards a more proactive and informed approach to FAIR data management
- > **roles** .. emerging roles .. chief data officer .. data steward .. organization members are sent to external workshops or training (e.g. GO FAIR Foundation) to lead the first internal projects
- > **processes** .. alignment discussions on metadata centralization .. growing commitment to embedding FAIR principles in data management practices.
- > **knowledge** .. people are sent to external workshops or training to understand FAIR .. there are pockets of FAIR knowledge (e.g. individuals) but no shared organizational knowledge .. some individuals connect with external communities of practice
- > **infrastructure** .. there is an understanding of the data challenges and a recognition of the importance of FAIR data.. initial plans for tooling and infrastructure for POCs start to take shape.. some centralization of metadata.. organization starts thinking about and implementing published metadata catalogs, controlled vocabularies

Dimension	Key notions
FAIR data	Data, metadata, data products
FAIR leadership	Types of leadership necessary for FAIR implementation
FAIR strategy	Approaches to implement FAIR data principles, use and business cases
FAIR roles	What kind of (human) roles are necessary to implement FAIR
FAIR Processes	Which processes must we explicitly implement
FAIR knowledge	What needs to be known for FAIR implementation
FAIR tools and infrastructures	From persistent identifiers to controlled vocabularies to semantic models

Table FAIR maturity matrix: Dimensions.

FAIR Maturity Matrix V1.0 © 2024 by [Pistoia Alliance](#) is licensed under. [CC BY 4.0](#)



Lopende projecten: Duaal PhD traject

- › PhD Project vanuit Rijksorganisatie voor Ontwikkeling, Digitalisering en Innovatie (ODI, BZK) i.s.m. LUMC (Leiden)
- › Focus op **semantische toegangscontrole**



Lopende projecten: Duaal PhD traject

- > PhD Project vanuit Rijksorganisatie voor Ontwikkeling, Digitalisering en Innovatie (ODI, BZK) i.s.m. LUMC (Leiden)
- > Focus op **semantische toegangscontrole**

- > Wie kan *toegang* krijgen tot welke **data**, op welk **moment** voor welk **doeleinde** voor **hoe lang**
- > Lezen, Schrijven, Aanpassen, ... ?

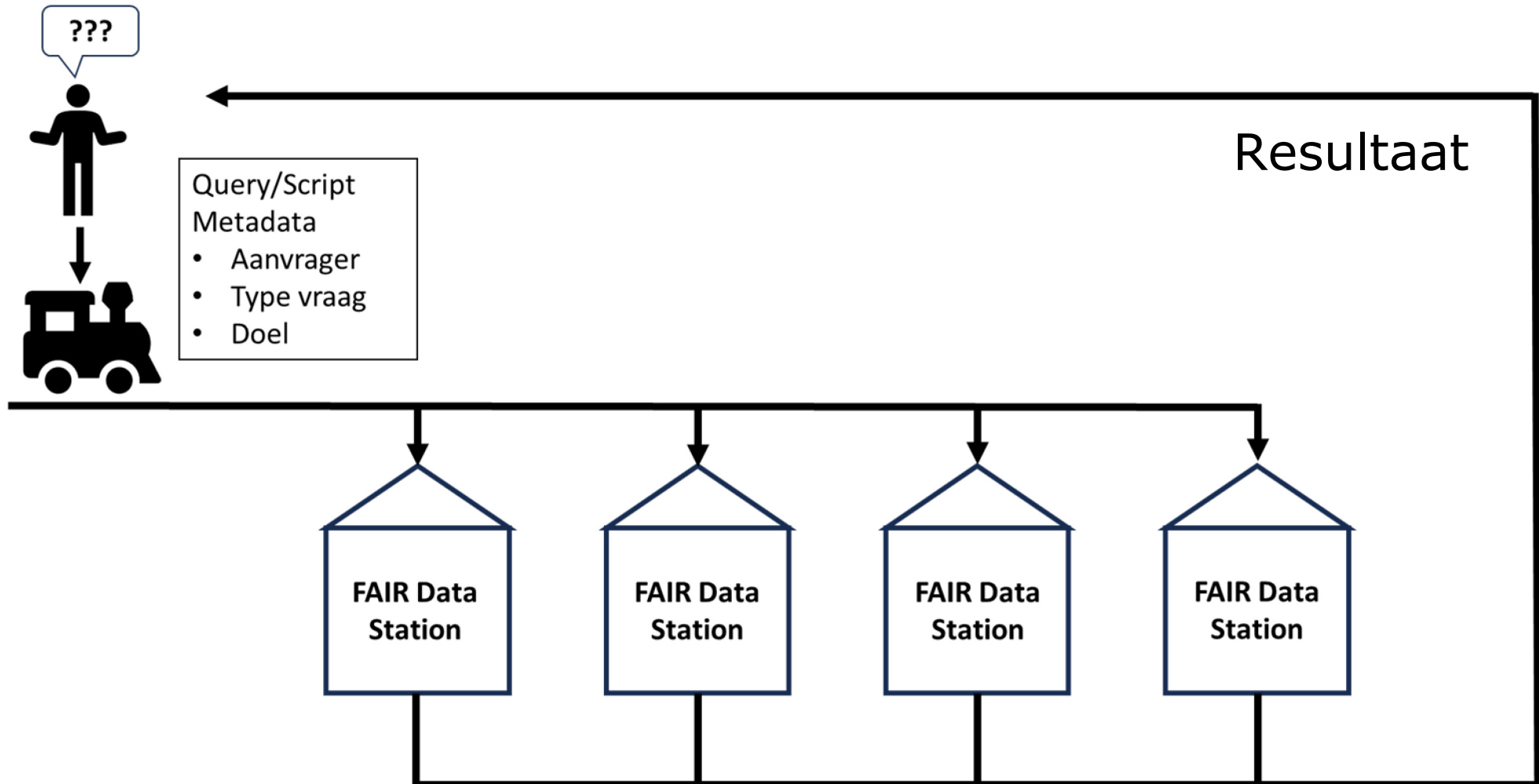


Waarom toegangscontrole?

- > FAIR Data $\neq$ Open Data
- > Veel redenen om informatie geheim te houden, bijvoorbeeld
 - Patiënten data
 - Confidentiële informatie (DepV, STG)
 - ...
- > In een verbonden wereld is het correct beheren van toegang tot data van cruciaal belang

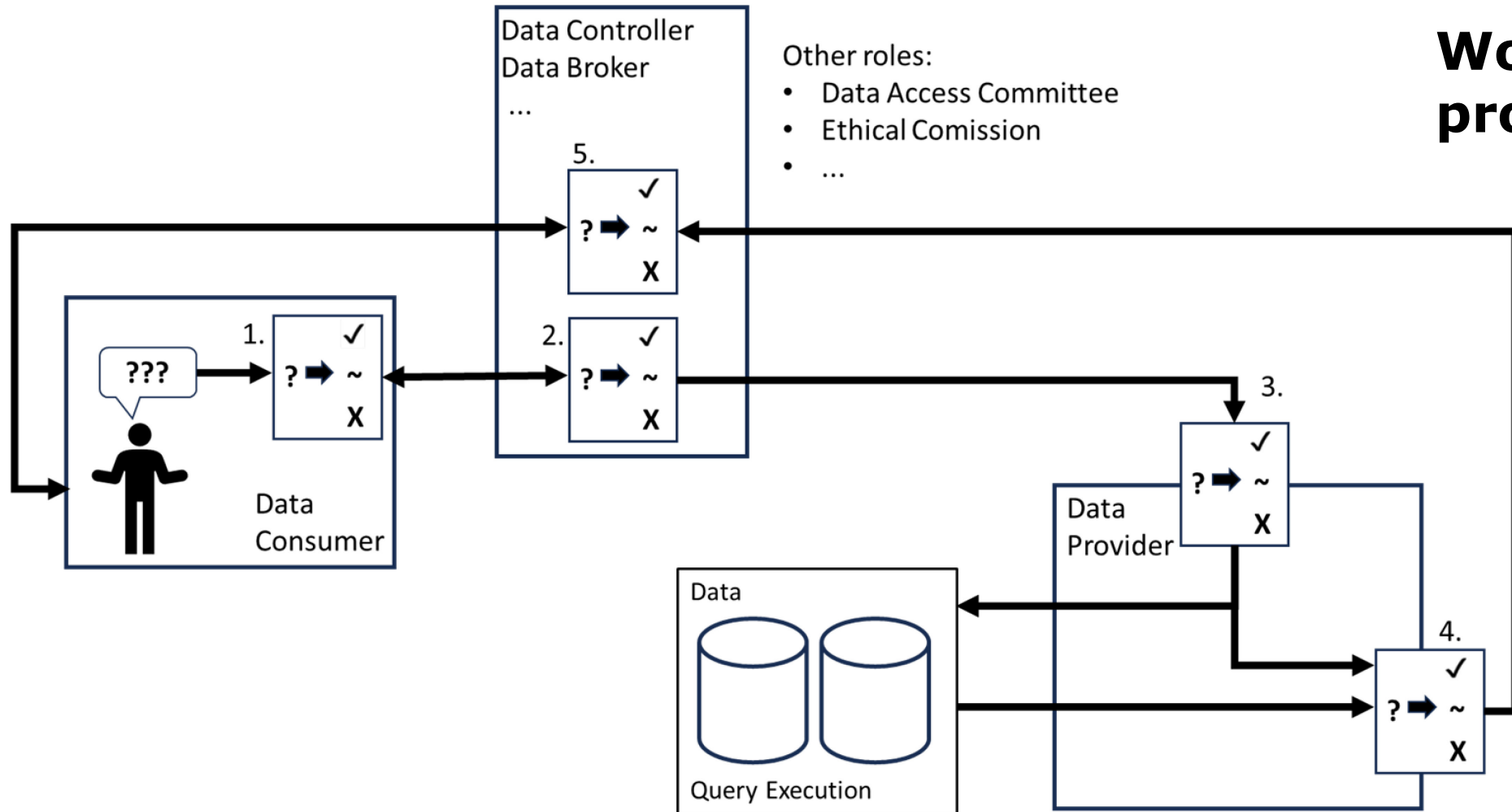


Personal Health Train





Federatieve toegangscontrole

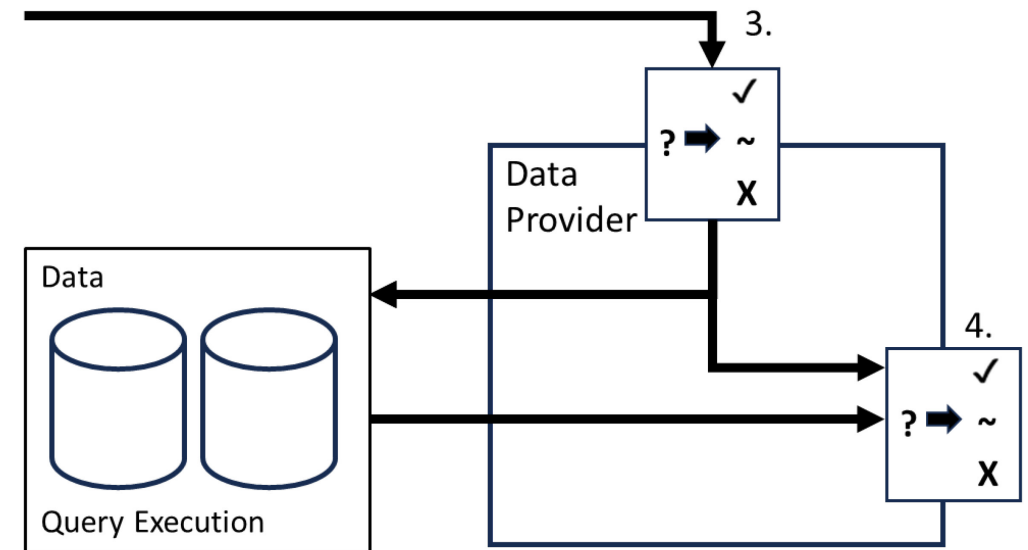


Work in progress!



Federatieve toegangscontrole

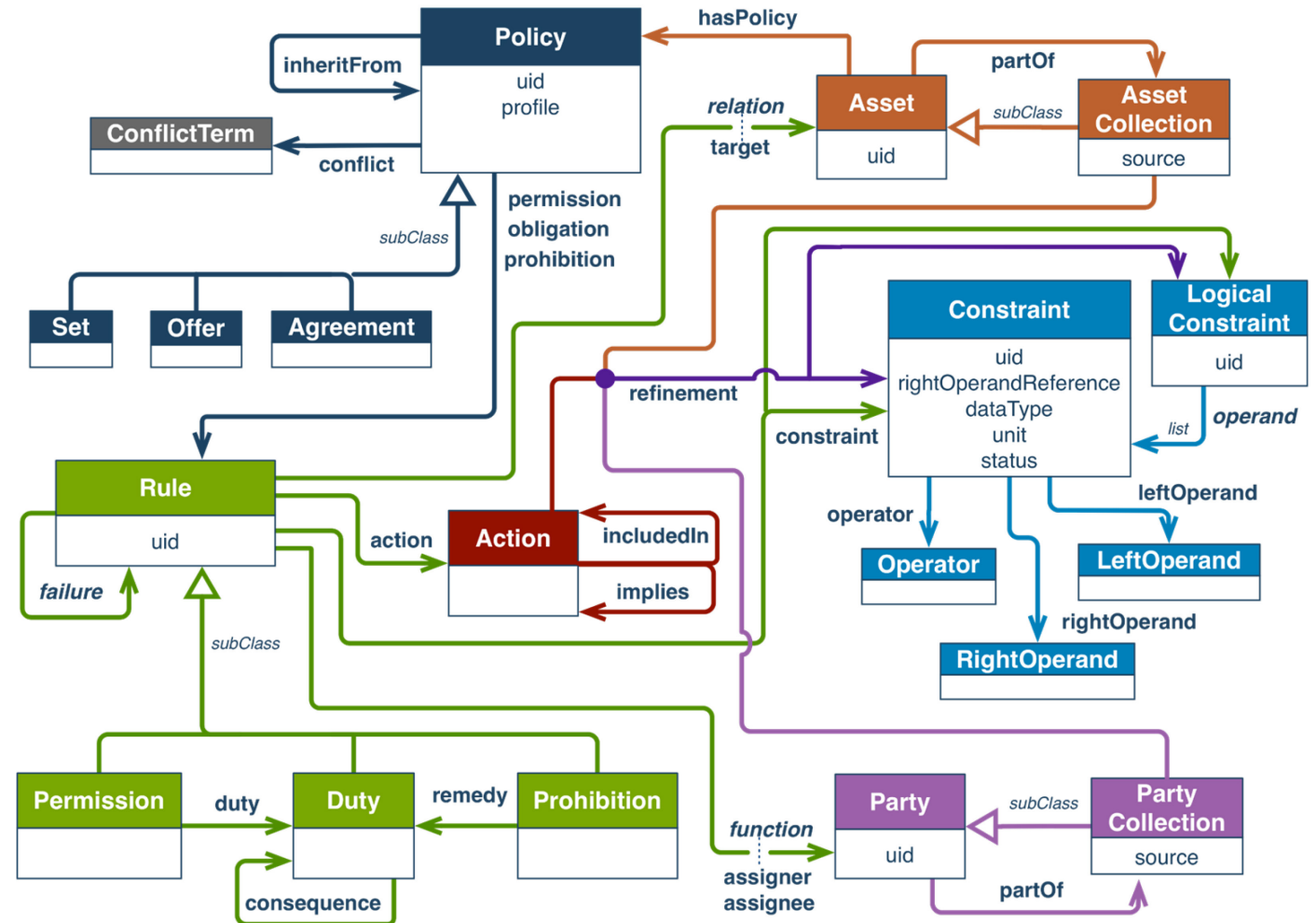
- > Focus op het niveau van een FAIR Data Station
- > Belangrijke aspecten (o.a.)
 - Definitie van data gebruiksregels
 - Implementatie en handhaving van business rules
 - Rol van een data steward





Implementaties

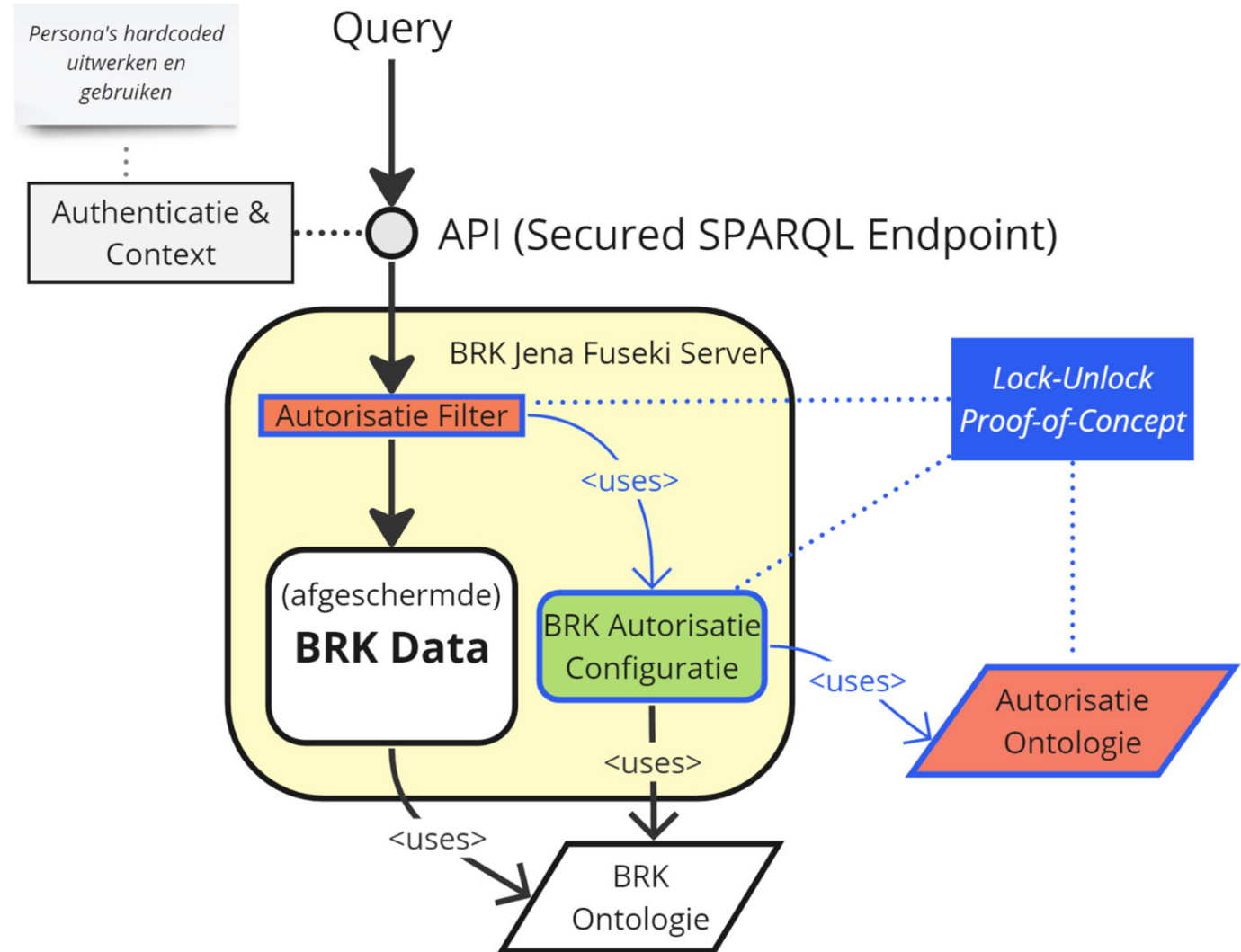
- > Open Digital Rights Language (ODRL)
- > V2.2 is een W3C recommendation
- > Informatiemodel om **usage licenses** mee te beschrijven
- > Behoeft een **evaluator** om toegang beslissingen te maken





Implementaties

- > Authorization Ontology door het Kadaster
- > Ontologie ontwikkelt specifiek voor toegangscontrole
- > Direct gekoppeld aan handhavingsmechanismen





Ontwikkelingen

- › Veel initiatieven voor semantische toegangsverlening
- › Federatieve Toegangsverlening - FDS & VNG
- › LockUnlock project - Kadaster
- › Internationaal
 - mu-authorization
 - ODRL
 - ...
- › Voor een waardevolle, federatieve FAIR implementatie moeten óók de toegangscontrole mechanismen **interoperabel** zijn



Nictiz

De Nederlandse kennisorganisatie voor digitale informatievoorziening in de zorg.

1. Ontwikkelen van een visie op het zorginformatiestelsel en architectuur
2. Ontwikkelen (met het veld) en beheren nationale standaarden (zibs, informatiestandaarden, terminologie)
3. Duiden, adviseren en kennis delen (richting VWS)
4. Participeren in nationale én internationale interoperabiliteitsprojecten (EHDS)
5. Stelselbeheerder van het Gezondheidsinformatiestelsel (Stelselregie samen met stelselhouder VWS)





18-10-2024



Onderzoek

Kwaliteit

Architectuur

Beleid

Ziekenhuis

GGZ

Verpleeghuis

Apotheek

Authenticatie

Autorisatie

ZIBs

Toestemming

Paramedisch

Mantelzorg

Burger

Identificatie

Kraam & jeugd

GYM

Adresboek

Huisarts

EUROPA

30

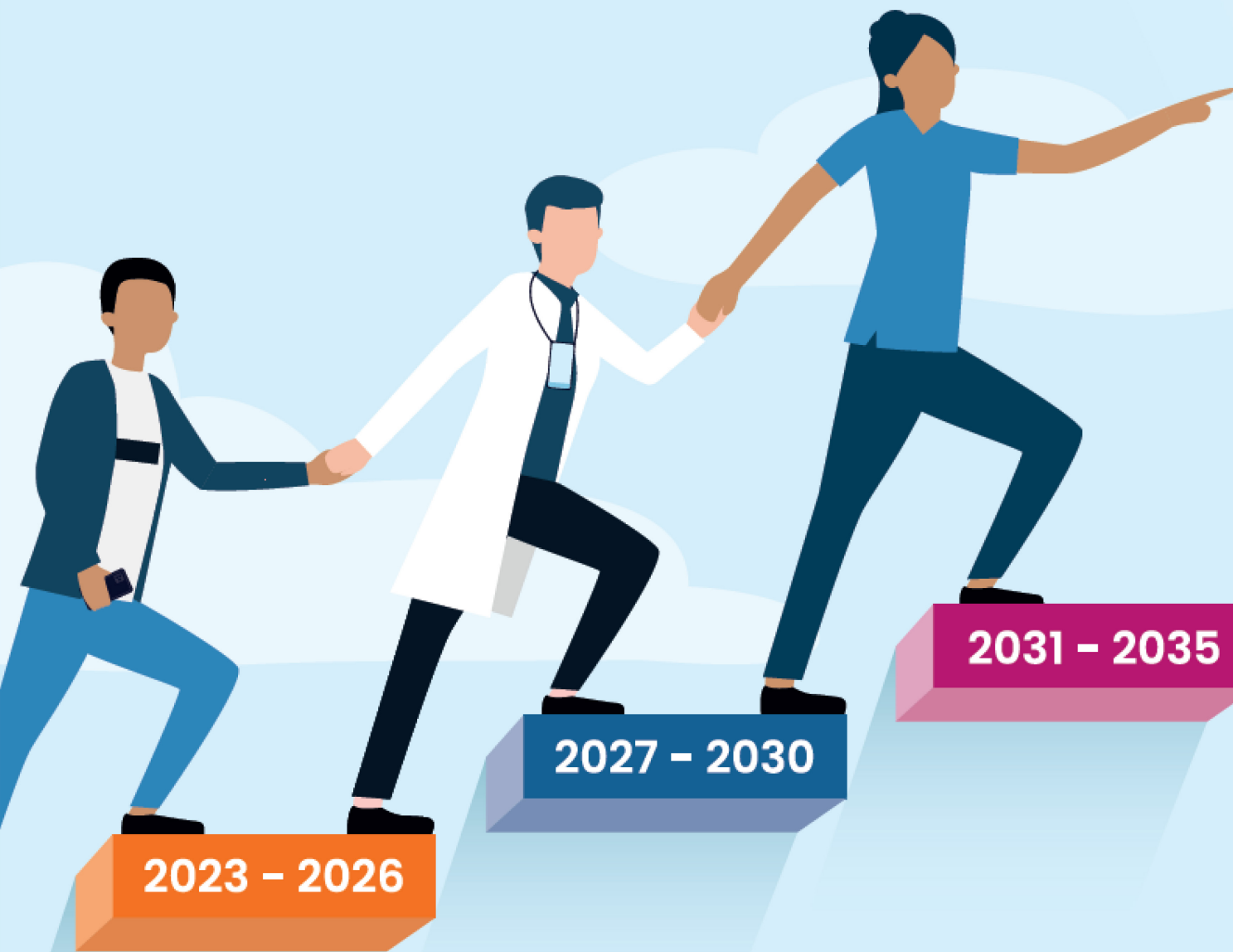
80

30

30

30

Nationale Visie en Strategie



Dromen

**Integraal
georganiseerd**

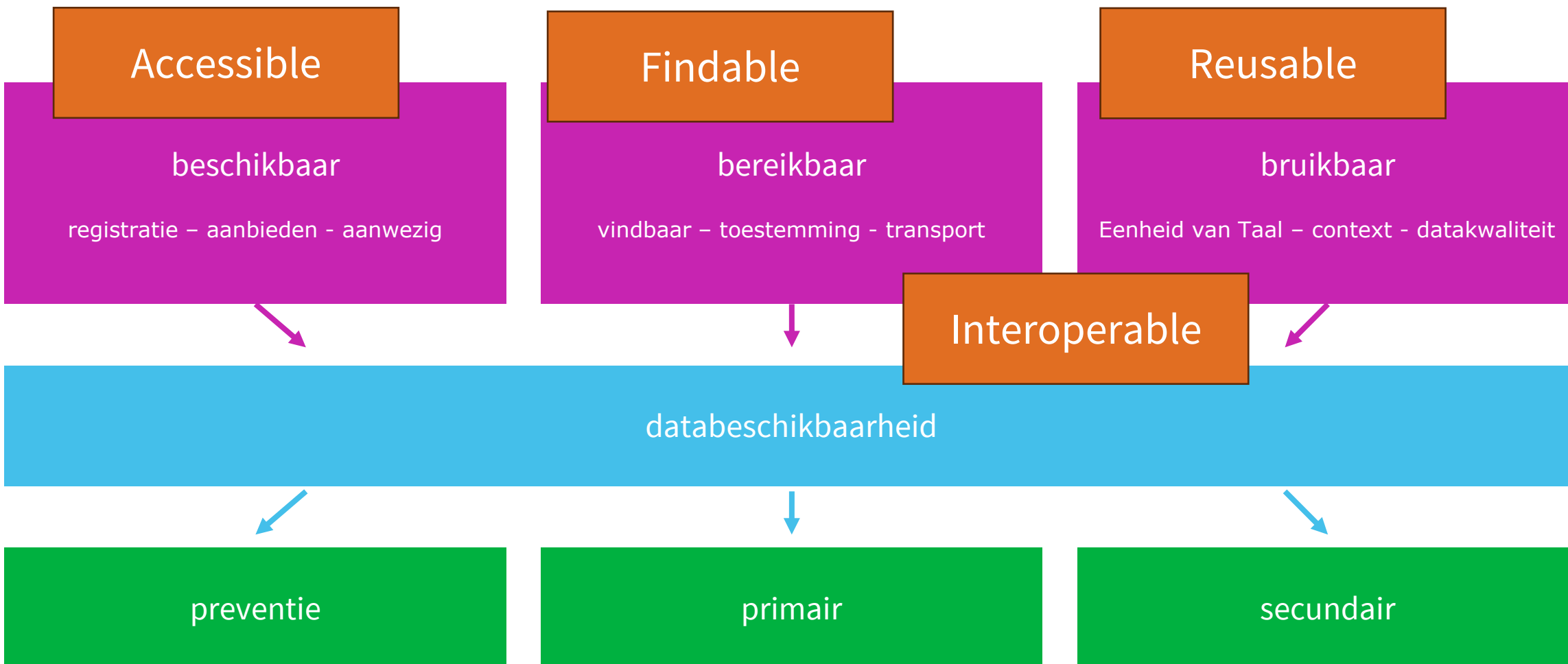
Denken

**Netwerk
georganiseerd**

Doen

**Interoperabiliteit
georganiseerd**

Databeschikbaarheid



Visie op het gezondheidsinformatiestelsel

Speelveldmodel

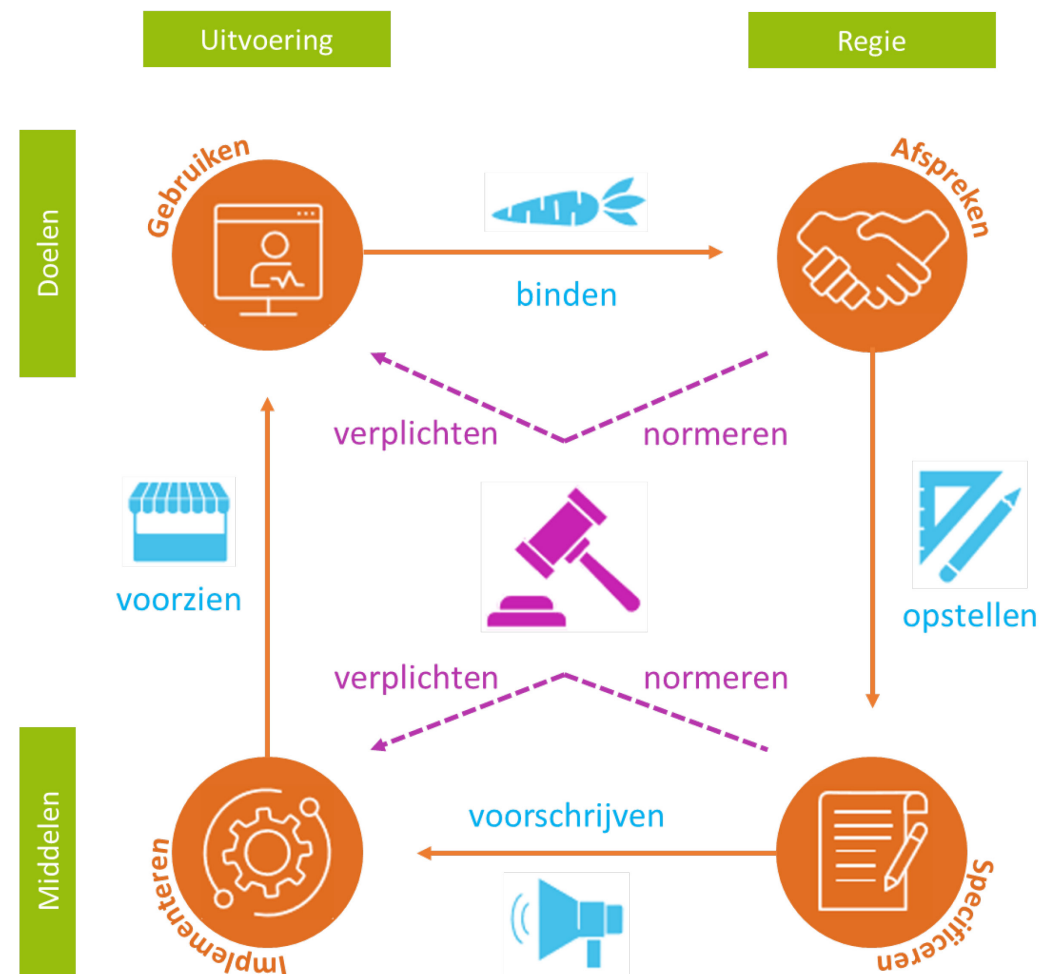
Nationale Visie en strategie:

- Gezamenlijk gebruik van informatie vraagt om gezamenlijke afspraken.

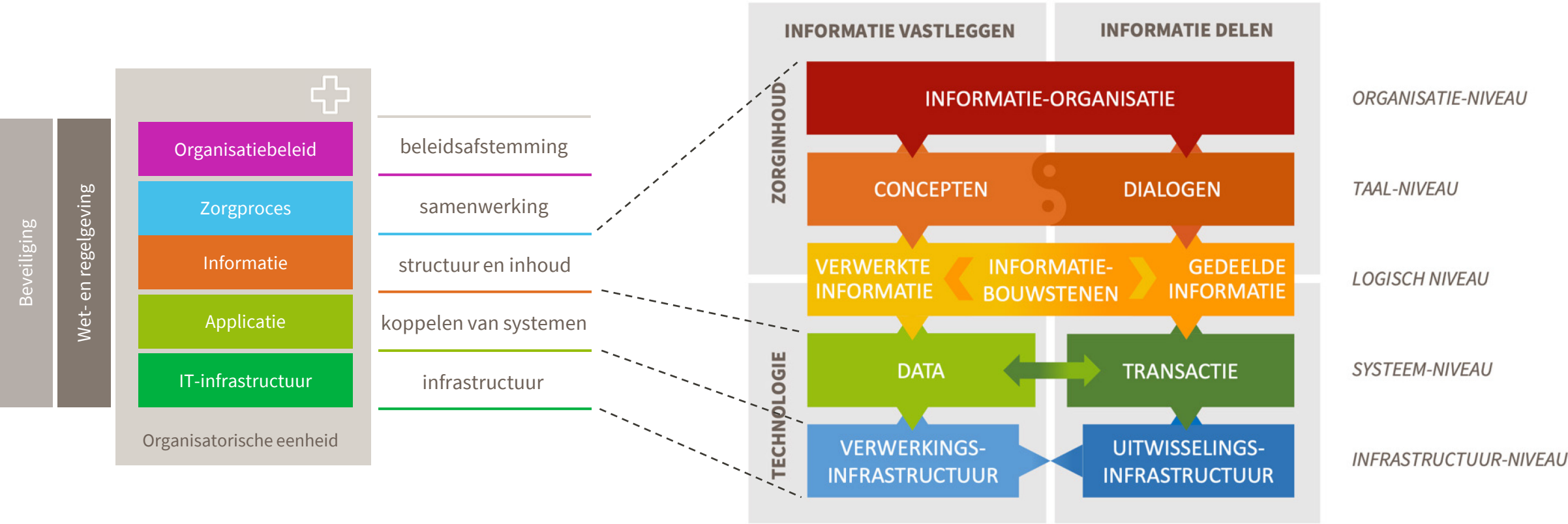
Hoofdactiviteiten nodig voor een werkend informatiestelsel (databeschikbaarheid):

- Gebruiken
- Afspreken
- Specificeren
- Implementeren

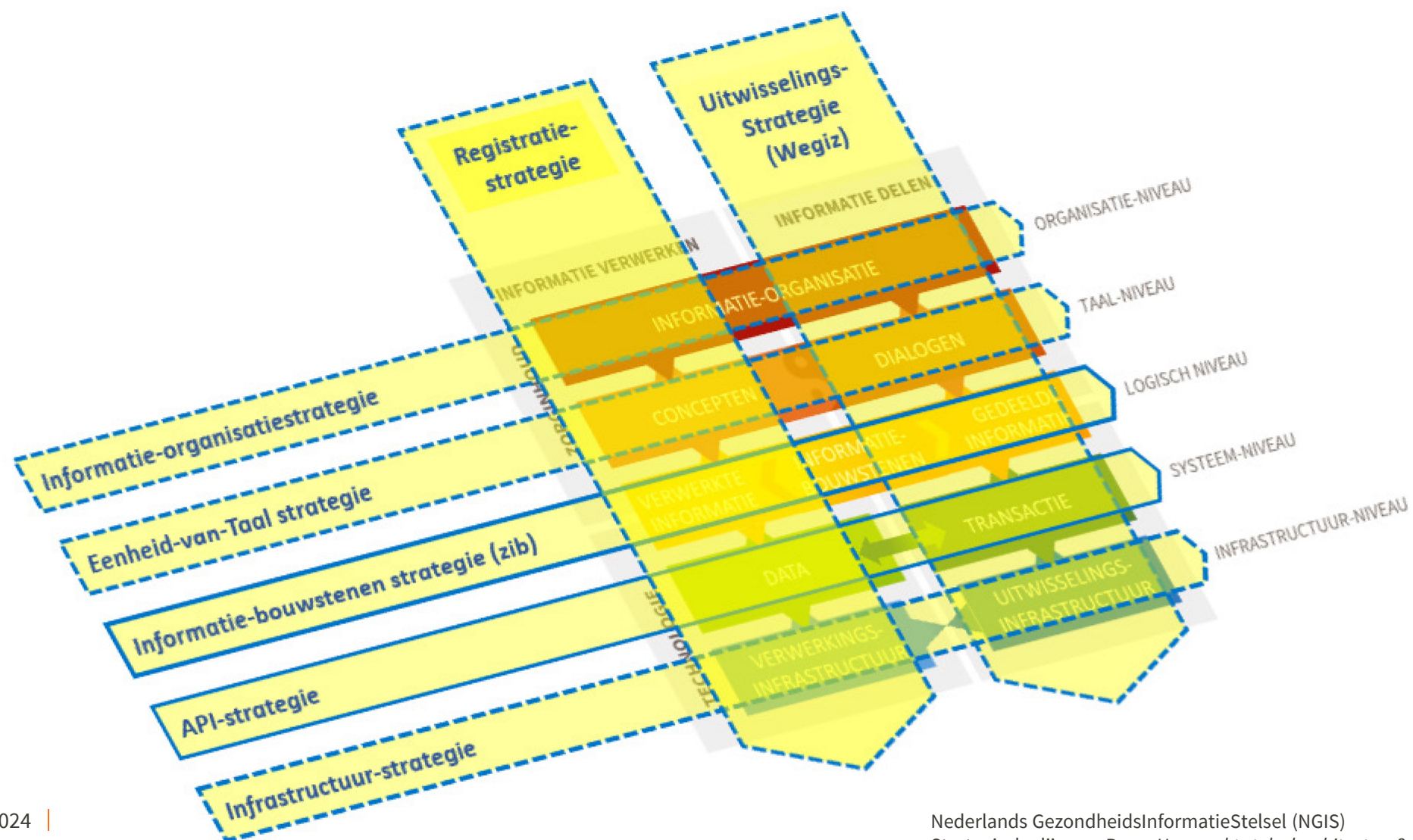
Primair en secundair datagebruik



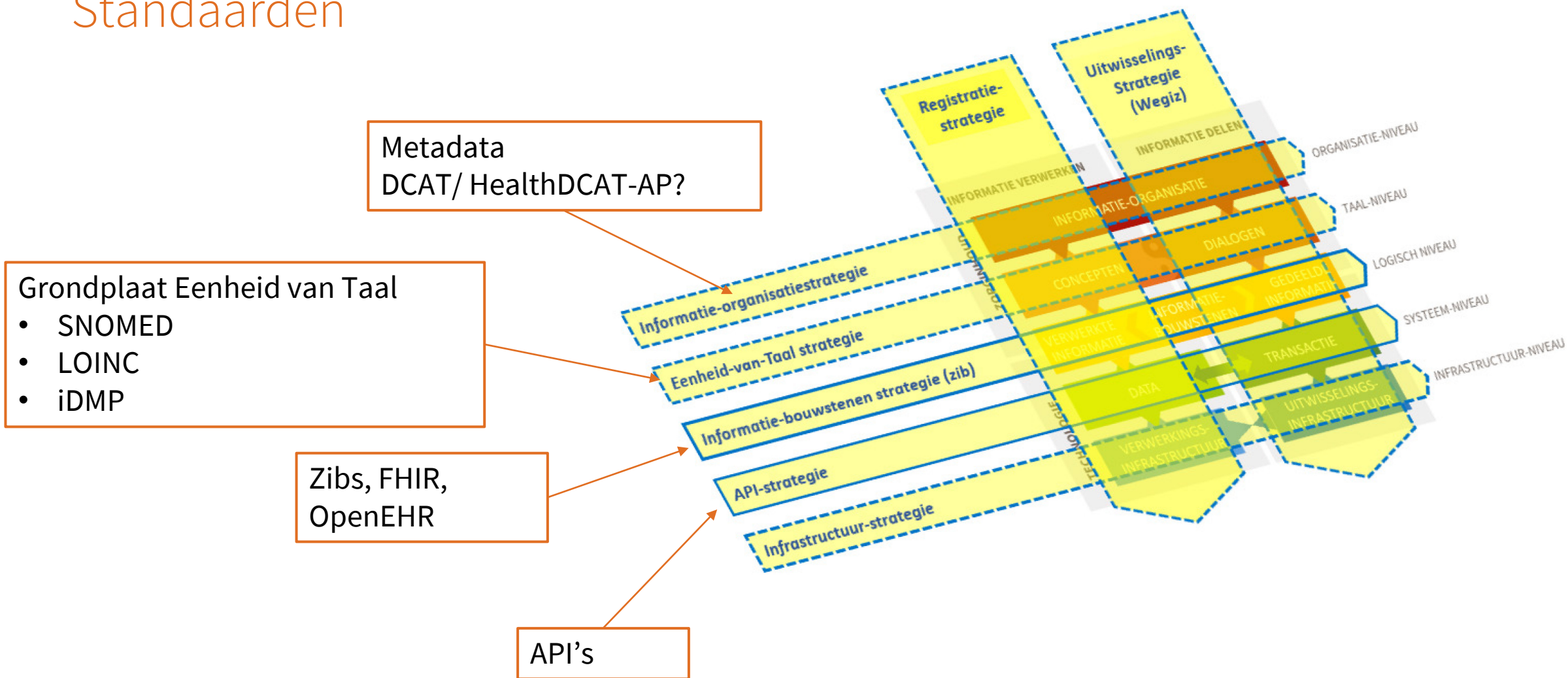
Interoperabiliteitsmodel



Welke afspraken zijn nodig in het gezondheidsinformatiestelsel?



Standaarden





FAIR is NOT a (next) standard

Barend Mons, October 2024



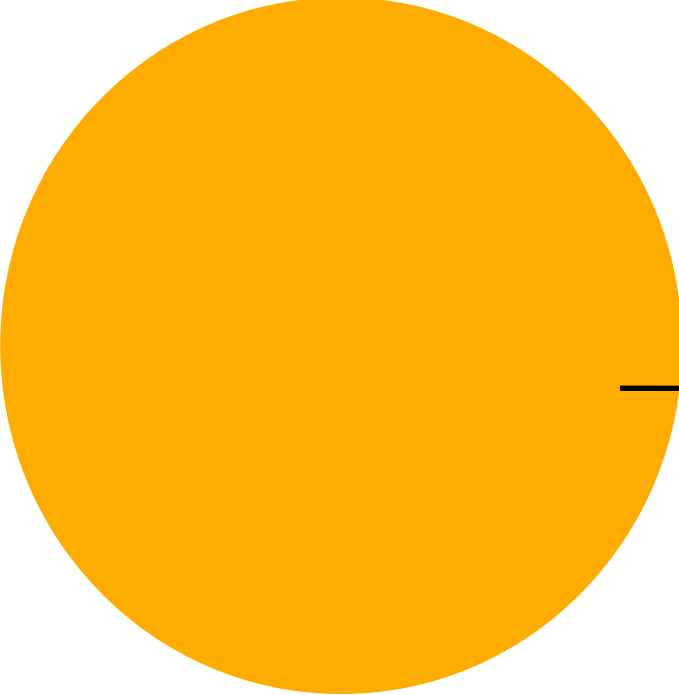
**Unity of Language < > Google
Translate**



Good standards should not be replaced



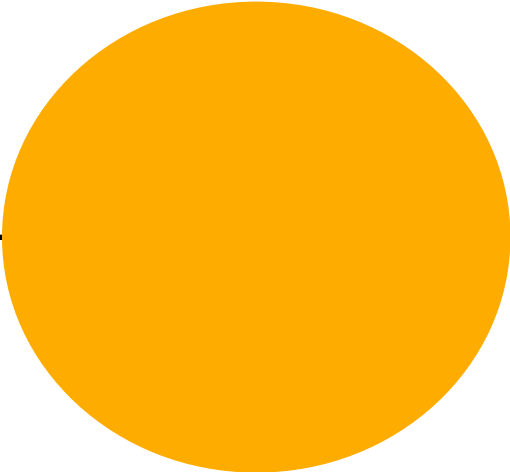
Recognized expert communities



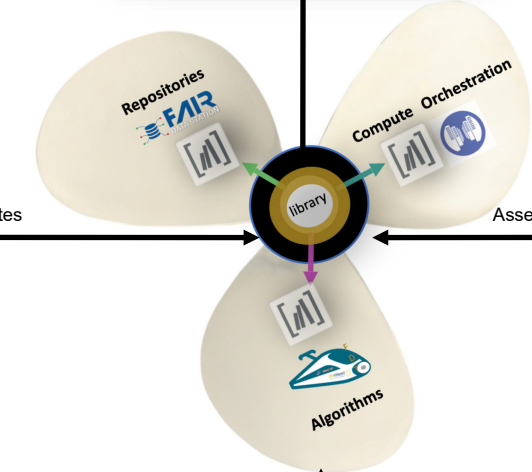
A global network of 'stations'



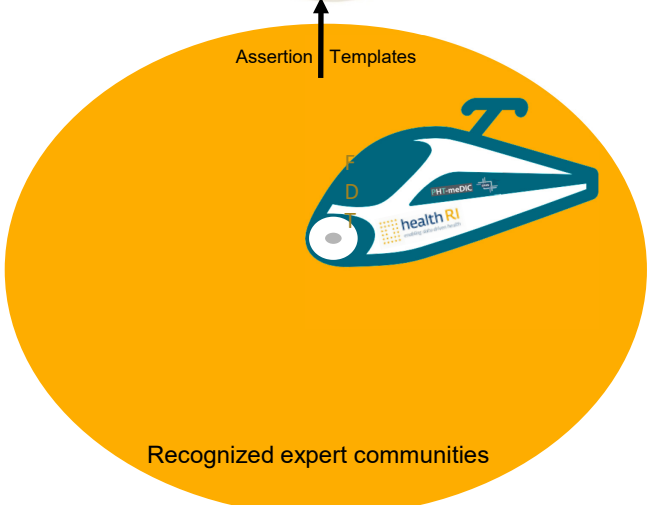
Recognized expert communities



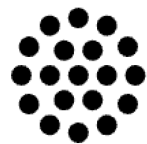
Specifications for orchestration infra



Assertion Templates



Recognized expert communities



International
Science Council



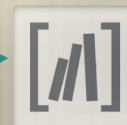
Repositories



CDIF



Compute Orchestration



brary



Algorithms



COMMITTEE ON DATA
CODATA
INTERNATIONAL
SCIENCE COUNCIL



GOSC
Global Open Science Cloud

130 semantic types (so far)

See also Semantic Type Workshop August 30 2023: <https://osf.io/ywneu/>

[illegible]



Content experts (with or without deep knowledge about FAIR)

< 30 min!

A screenshot of a survey interface. At the top, the word "yes" is typed into a text input field. To the right of the text field is a green dropdown menu with a downward arrow, also displaying the word "yes". Below the text field, there is a list of three options: "yes" in a green circle, "no" in a red circle, and "Not yet sure" in a yellow rounded rectangle. At the bottom right of the form, there is a small icon of a pencil.

Publish your choices

[SEARCH](#)[FAIR SUPPORTING RESOURCES](#)[DASHBOARD](#)[ABOUT](#)

FAIR Connect is an Open Access environment promoting the collaborative reuse of FAIR Supporting Resources and the dissemination of good practices for professional FAIR-Data stewardship.

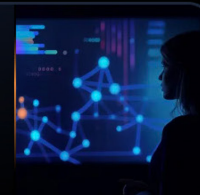
SEARCH FAIR SUPPORTING RESOURCES

FAIR metadata for FAIR Supporting Resources →



FAIR SUPPORTING RESOURCES DASHBOARD

Get insight on nanopublications →



SUBMIT AN ARTICLE

Submit a short-form article of 2000 words or less →



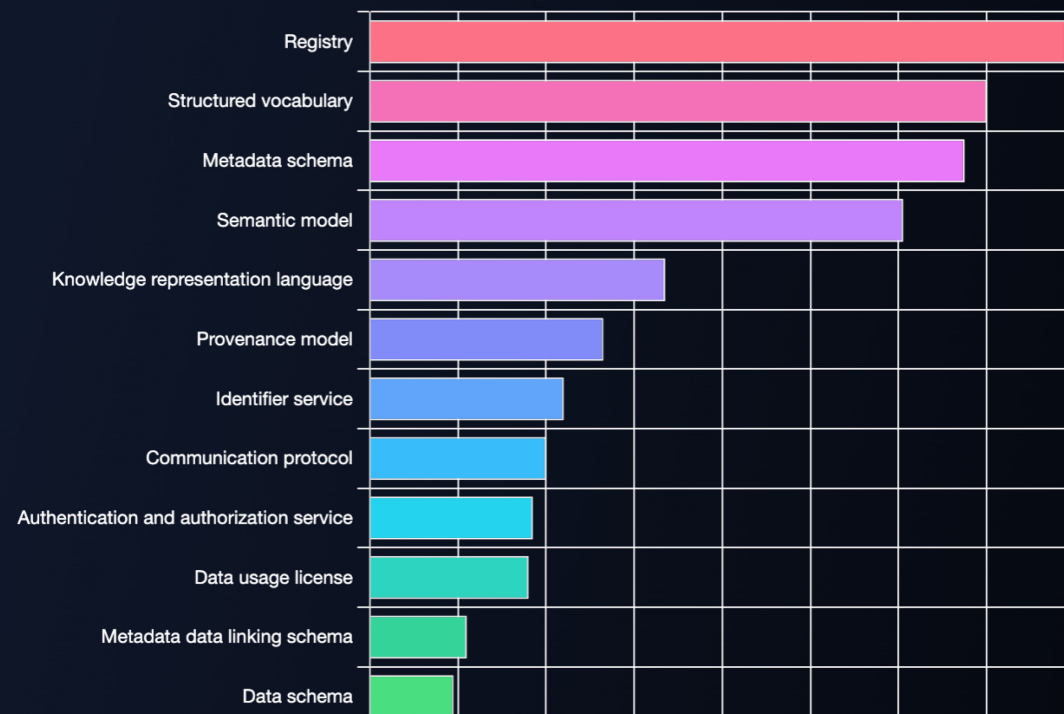


Dashboard

997 Published FSRs

Click on the bars to see a list of FSRs labelled by the selected Type

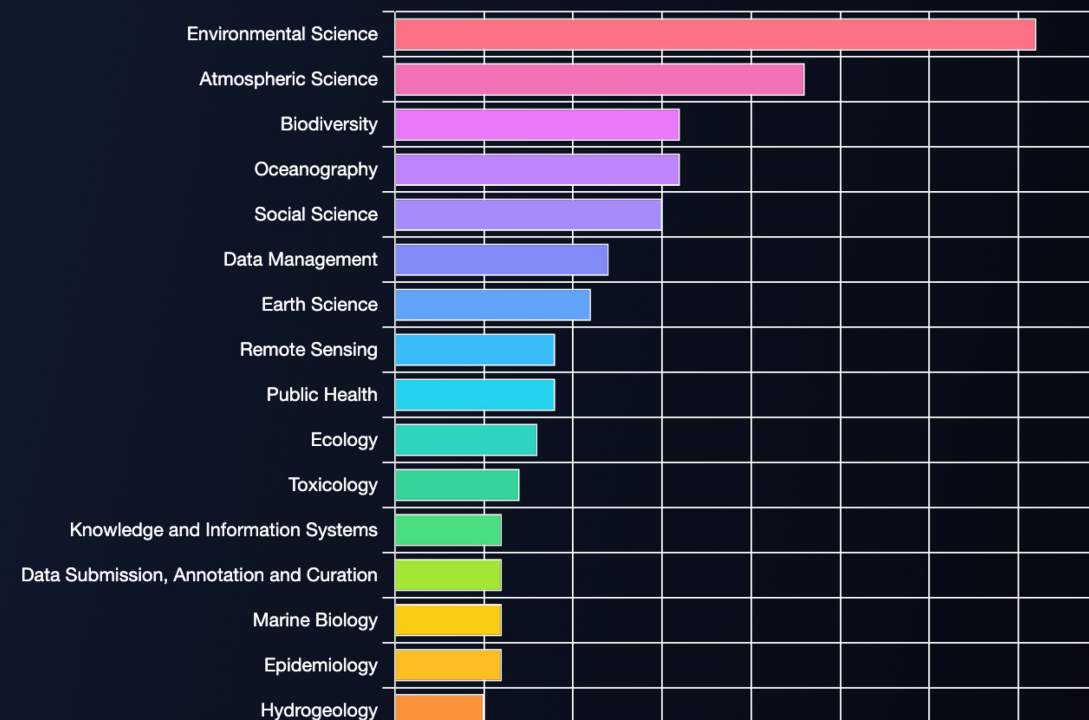
Sort by: ☒ Number of FSRs ☐ A-Z



353 Published FIPs

Click on the bars to see FIPs that are used by the community in a Domain

Sort by: ☒ Number of FIPs ☐ A-Z





Under Design



FAIR enabling resources for this semantic type

Wizard

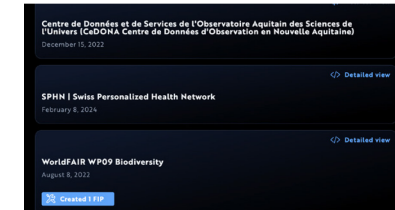
yes | yes

yes

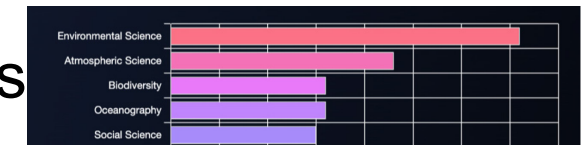
no

Not yet sure

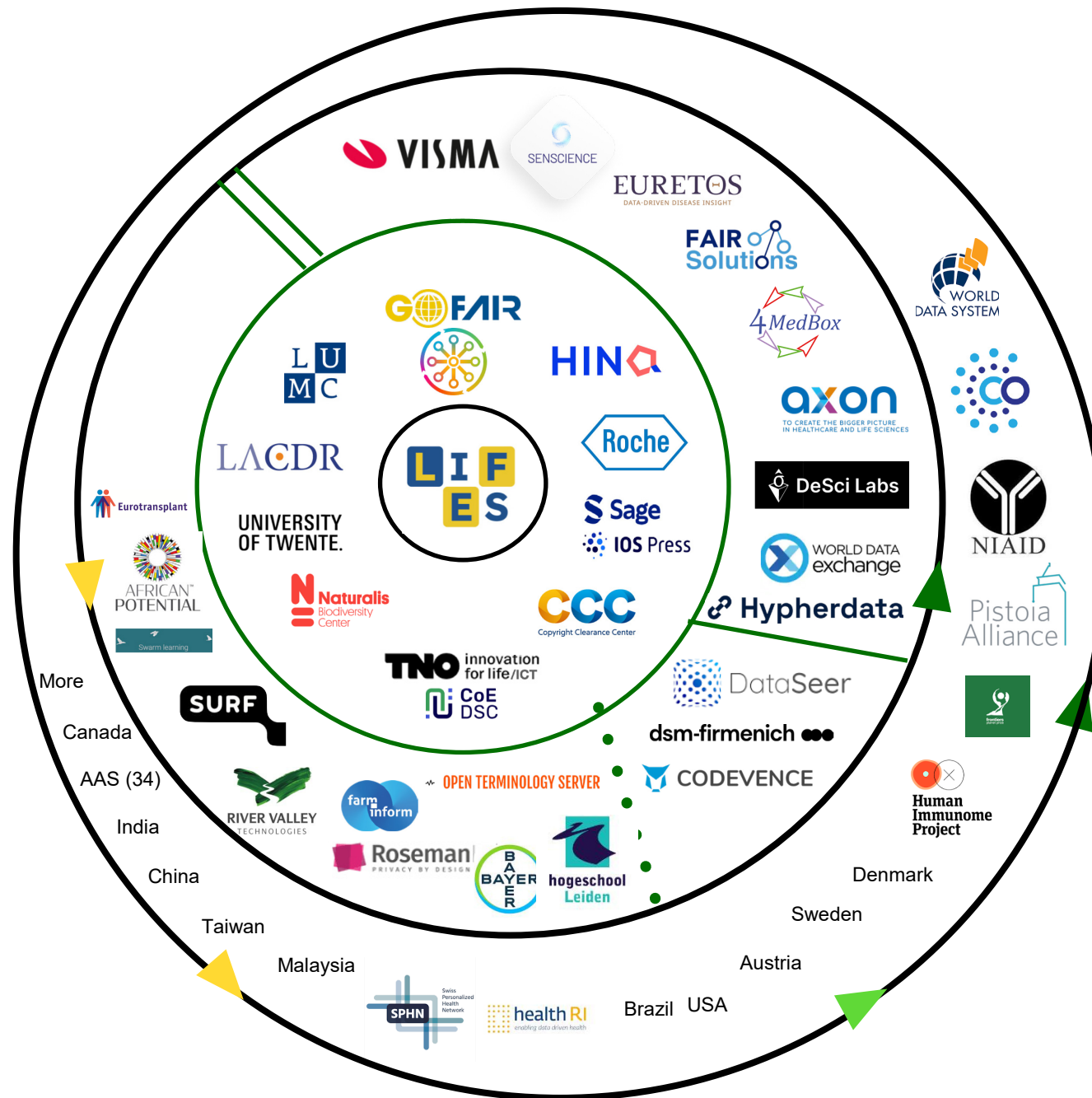
Similar communities



Similar (reusable) FIPs



Convergence by Convenience



More

Canada

AAS (34)

India

China

Taiwan

Malaysia

Brazil USA

Austria

Sweden

Denmark



Literatuur

Bonino da Silva Santos, L.O., Ferreira Pires, L., Graciano Martinez, V. et al. Personal Health Train Architecture with Dynamic Cloud Staging. SN COMPUT. SCI. 4, 14 (2023). <https://doi.org/10.1007/s42979-022-01422-4>

Innovalor in opdracht van Forum Standaardisatie, “[Onderzoek naar de stimulering van FAIR data principes bij de overheid](#)”, 2021

Jacobsen, Annika, et al. “FAIR Principles: Interpretations and Implementation Considerations.” *Data Intelligence*, vol. 2, no. 1–2, 2020, pp. 10–29, https://doi.org/10.1162/dint_r_00024.

PBLQ in opdracht van Rijksprogramma voor Duurzaam Digitale Informatie Huishouding, [Duurzaam Digitaal Databeheer Bij de Rijksoverheid: Een Verkenning](#). p.23, 2021.

Wilkinson, Mark D., et al. “The FAIR Guiding Principles for Scientific Data Management and Stewardship.” *Scientific Data*, vol. 3, no. 1, Mar. 2016, p. 160018, <https://doi.org/10.1038/sdata.2016.18>.



Reserve: FAIR data versus Linked data

Overeenkomsten

- › Beide hebben als doel om gegevens machine leesbaar en herbruikbaar te maken, maar FAIR legt in bredere zin de nadruk op toegankelijkheid en de rijkdom van metadata.
- › Beide bevorderen interoperabiliteit, maar FAIR richt zich op richtlijnen op hoog niveau, terwijl Linked Data zich concreet en specifiek richt op RDF.
- › Beide hechten waarde aan de rol die metadata spelen bij het beter ontdekken en hergebruiken van data.

Verschillen

- › FAIR-principes zijn algemener en hebben betrekking op de algemene toegankelijkheid en herbruikbaarheid van gegevens, terwijl Linked Data een concreet technisch raamwerk is voor het structureren en verbinden van data.
- › Linked Data gaat over specifiek over integratie en het koppelen van gegevens uit verschillende bronnen, ("inrichten") terwijl FAIR zich meer algemeen richt op vindbaarheid, toegankelijkheid, interoperabiliteit en hergebruik te verbeteren. ("richten")
- › Linked Data is afhankelijk van specifieke formaten en technologieën, terwijl FAIR ruimte laat voor het gebruik van verschillende technologieën zolang ze maar voldoen aan relevante community-standaarden.

Samenvattend:

- FAIR Data en Linked Data overlappen elkaar in hun focus op interoperabiliteit en het machinaal leesbaar maken van data. Ze verschillen echter in aard en scope, principes ("richten") versus specificaties ("inrichten"), specifieke technische implementaties en algemene doelstellingen.
- FAIR-principes zijn breed, terwijl Linked Data een concreet en specifiek raamwerk betreft.