



Open Consultation Architecture design v1.0

June 6th 2023

Recording

This meeting will be recorded

Transcription
Video



Objective

Feedback – Optimization

Support base



Agenda



- House rules
- Short introduction Health-RI
- Health data infrastructure for research and innovation
- Break (+/- 10:30 am)
- Storyline: Research and innovation
- Storyline: Making data available
- Next steps

House rules



Microphone off

Moderator turns microphone on/off for you



No questions during the presentation

Time for questions from the Chat after each block



Post questions in the Chat

We will cover as much questions as possible today

Remaining questions will be answered later by e-mail

Please clearly indicate who is asking the question



We do not work with “Raise hand” function

Short introduction Health-RI



- Mission
- Key message
- Obstacle Removal Trajectory (OVT)
- Stakeholders
- Nodes, PoCs and Use Cases

Mission

Better health(care) for
citizens and patients

by

reusing health data

with

an integrated health
data infrastructure

for

research and innovation





The Challenge?
Fragmentation!

Key message

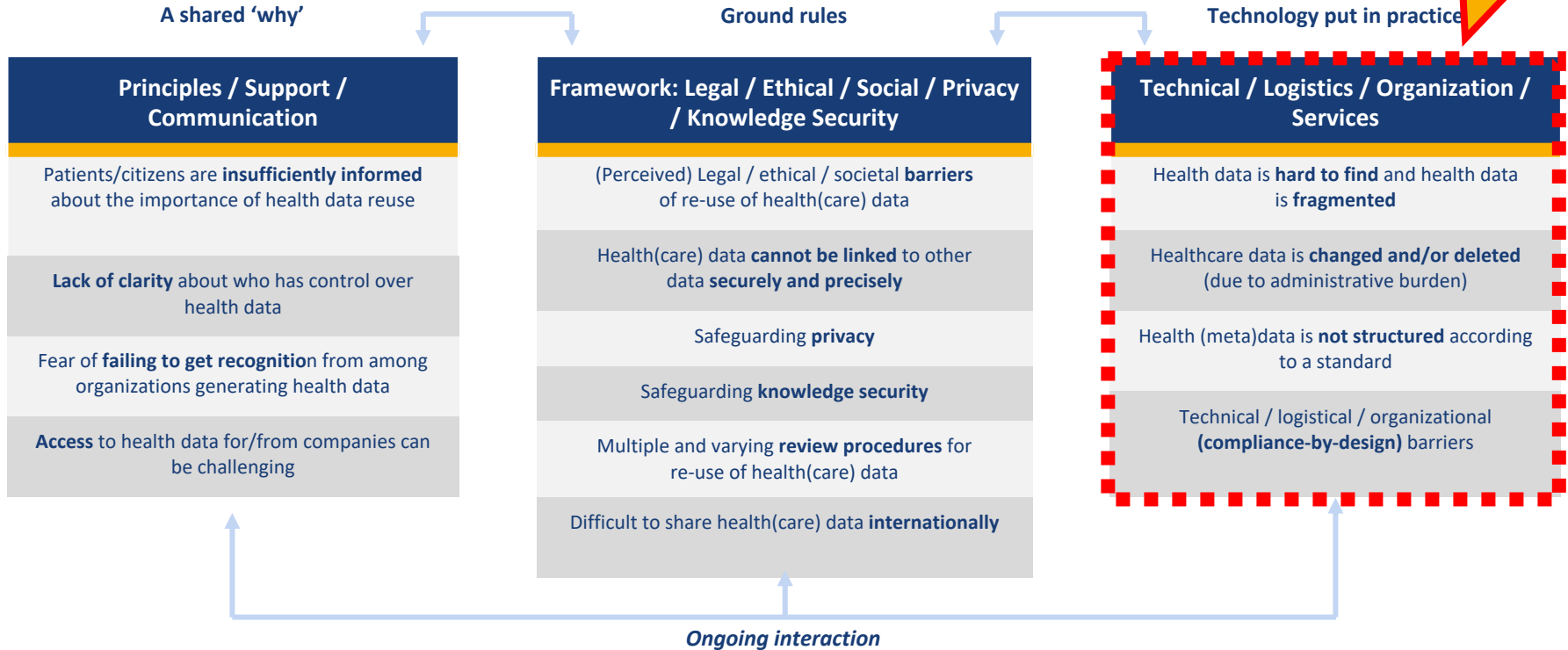
To achieve this mission, a number of preconditions are important:

- The infrastructures for care, research and innovation must be linked together and should be interoperable and integrated. The principles and processes for capturing and accessing data from the healthcare chain should be jointly designed.
- The necessity of integral sectoral legislation and regulations for health research must be taken into account.
- Needs national governance and procedures
- Setting up and maintaining a national health data infrastructure requires a utility service: a service of general economic interest that facilitates, leads and coordinates.

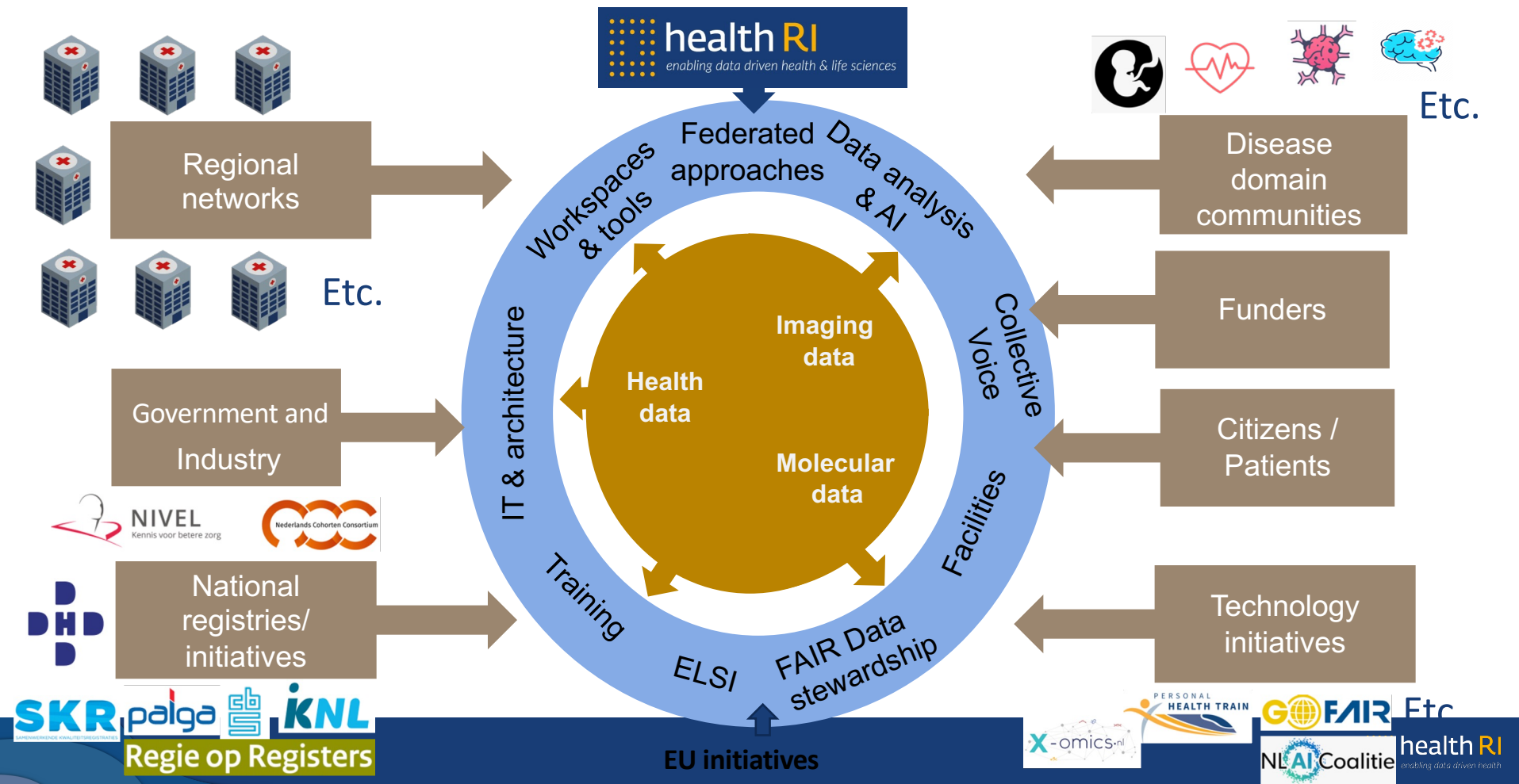
Obstacle Removal Trajectory (OVT)

14 obstacles in 3 clusters

Architecture:
Cluster 3
Solution 6, 7 en 8



Health-RI: integrated health data infrastructure for research and innovation



“Think big – Start small – Act now”



Think big:
integrated health data infrastructure



Start small:
PoCs, use cases



Short introduction Health-RI



Questions?

Health data infrastructure for research and innovation



- Joint architecture model
- Health data infrastructure for research and innovation
- Received comments and questions
- Your questions

Common care architecture model

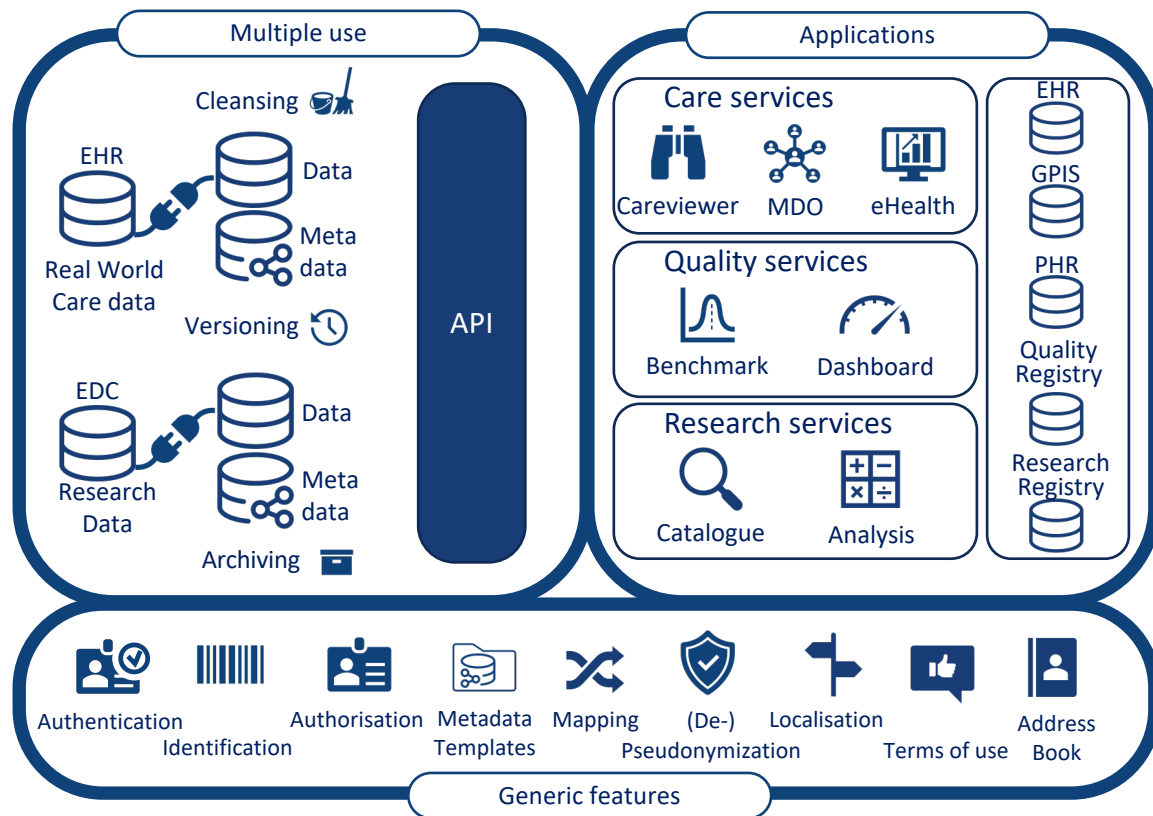


- Broad Support:
 - VWS, ZN, VZVZ, Nictiz, IHE, NFU/ AcZie, TWIIN and Health-RI
- Narrative:

“Each relevant health data point and the associated context are available to each authorized person within conditions of use”

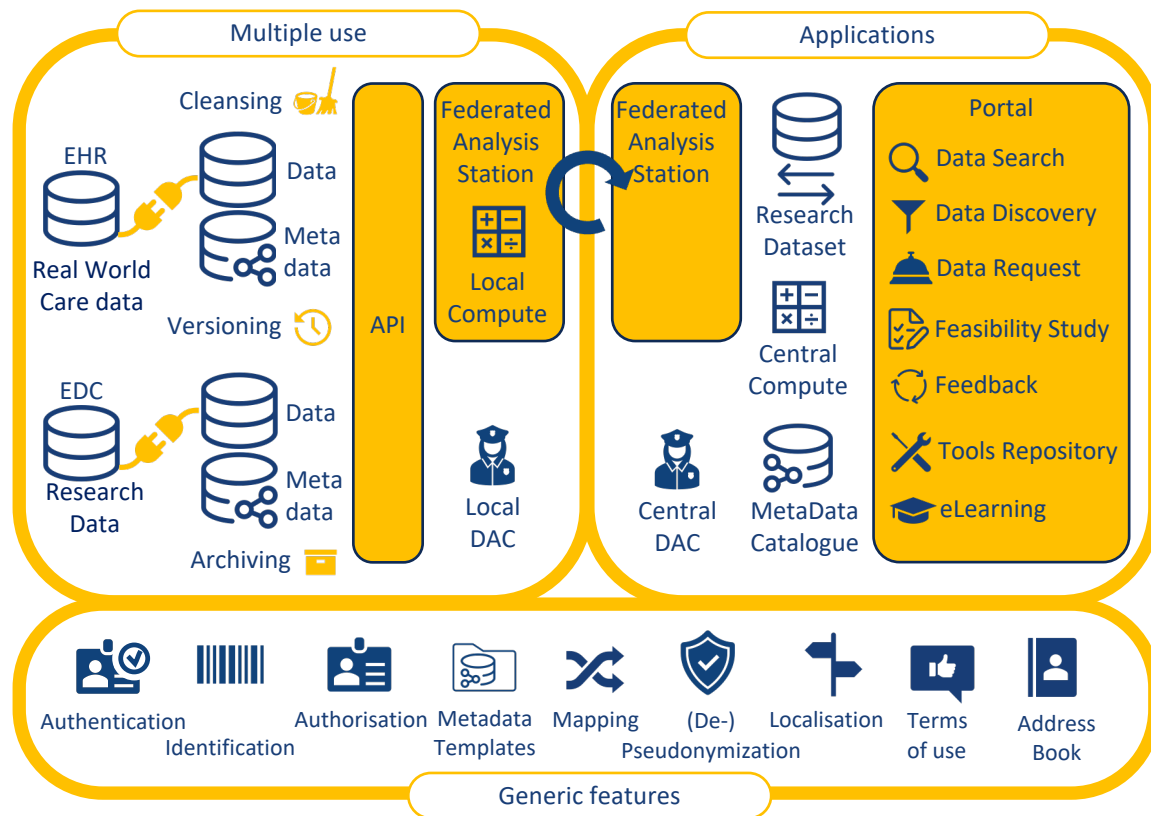
Generic model: 3 components

- Data centric
 - ✓ Separating Data from Functionality
- Unity of language
 - ✓ [RIVM Rapport 2018-0081](#)
 - ✓ Semantic interoperability
- Standard API
 - ✓ Standard Functions
 - ✓ Standardized Datasets
 - ✓ Standardized method Metadata (DCAT)
- Duplicate as little as possible
- Generic features
 - ✓ Based on international standards
 - ✓ No differentiation
 - ✓ Same for Primary & Secondary use



Health data infrastructure for research and innovation

- Portal
 - ✓ Research Ecosystem
 - ✓ Facilitate re-usable results
 - ✓ Broad toolset
- Metadata catalogue
 - ✓ DCAT
- Central analysis
 - ✓ Secure analyse/process environment
 - ✓ Central compute
- Federated analysis
 - ✓ Local compute
- Data Access Comity
 - ✓ Central
 - ✓ Local



Comments and questions received Architectural model (1/2)

Question / Comment	Reaction
PGO (Pesroonlijke GezondheidsOmgeving) (UK->PHR: Personal Health Record) as source/ Civilian as actor is missing	No room to name all the different sources and participants. Are included in architecture. See Roles and Participants
It is suggested that the H-RI ecosystem also provides the 'infrastructure' for primary use of health data and for making primary health data available for derived use.	That is not the intention. The H-RI Ecosystem benefits from useful care data for research and innovation and works closely with parties from the healthcare sector (e.g. CumuluZ), so that they make this data suitable for derived use.
Is it an idea to position the HRI architecture within existing reference architectures from the start?	There are several reference architectures, which we have studied, and to which we want to connect (see Architecture - Relevant initiatives in healthcare, research and personal domain) Suggestions are welcome.
The persistent data platform is not clear whether the intention is that this is foreseen as centralized/pooled persistence and/or whether a federated persistence is pursued.	We now work with 8 regional nodes, with a zoned persistent data platform as a shared service centre for the regional care parties. Parties/ systems can themselves offer persistent data to the H-RI ecosystem by means of standardized interface.
Shared service persistent data platform merges source data from multiple institutions that will be issued for as yet unknown research questions. This still requires organizational measures (e.g. setting up a foundation per data collection) and possibly verification of the lawful basis to do so.	Agree

Comments and questions received Architectural model 2/2

Question / Comment	Reaction
Who defines the API, who develops the API? Can it then be used by all regional nodes and used to access the regional data?	Questions that we will work out together. The idea is to define it together with the nodes. Perhaps we can turn this into a standard SAAS solution, which can be applied by the nodes
More explicitly distinguish between applications that: • Reusing single data, read: reusing individual patient records • Multiple data, e.g. for entire cohorts and populations	In the API we will standardize the different allowed actions on the data. We will work these out in more detail in later versions. In this we should pay attention to requesting information about 1 person, or about a group of people
On the left side of the picture, data from care is made suitable for research: cleansing, adjusting, deduplicating, enrichment, derived data. This data has been adjusted/filled in by researchers, and cannot be called a care file in this form.	On the left, this data has not been adjusted/filled in by researchers, and cannot be called a care file in this form. On the right, researcher could do an analysis processing environment in this area, where feedback can be given to source; This feedback loop will be described in more detail in a future version.
The (regional) nodes also provide the module application areas. If all data is located at the nodes, the module application areas can be controlled centrally, but it is not necessary to offer it centrally.	Agree. The mechanism by which the Health-RI hub application module finds or applies data can also be applied in a regional node. With which data from other nodes can also be found and applied there. It can be cost-effective to use the Health-RI hub scope module, instead of setting up your own.

Health data infrastructure for research and innovation



Questions?

(bio)break



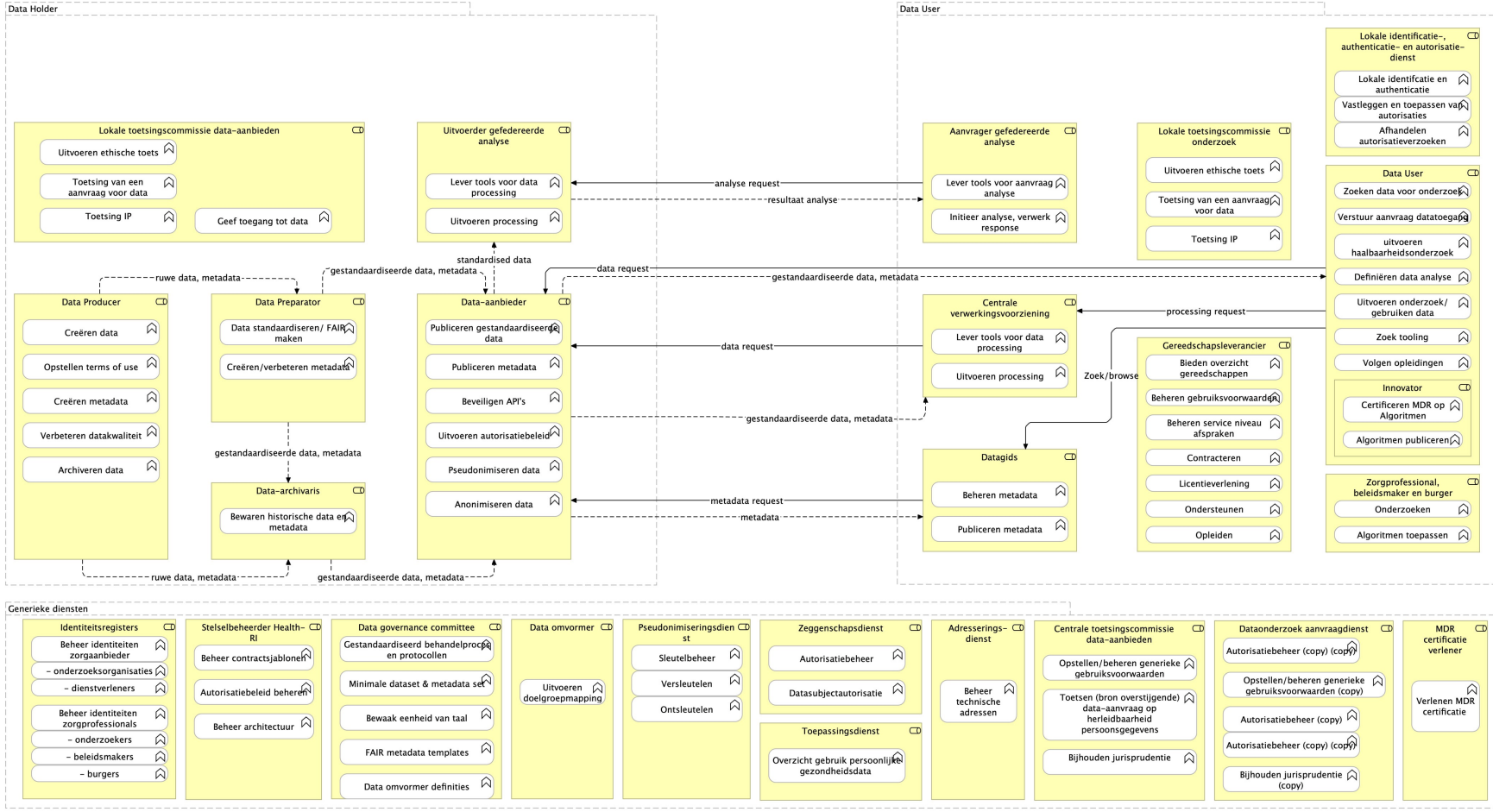
- 10:35 am we'll continue

Research and innovation storyline



- Roles
- Different steps from storyline
- Comments and questions received
- Your questions

Roles



Research and innovation storyline

- This overarching storyline provides functional insight into the research and innovation process to be supported.
- This leads to the following steps in the research and innovation lifecycle (based on Business Processes Research Data Cycle - HORA2 wiki):

Step 1. Research and innovation storyline

- **Researcher** wants to start a new study.
- The organization for which **Researcher** works, carries out a medical ethics review from the role of a **Local review committee launch research** and, if relevant, the Technology Transfer Office (TTO) is also consulted.
- **Research data request service** facilitates and administers process for data user to ethically and legally test the research at the **Local review committee launch research**.

Step 2. Research and innovation storyline

- **Researcher** designs and manages a data management plan from the start, with which the research data and results are made suitable for multiple use.
- **Researcher** uses code/ modelling pages and metadata templates prescribed by the Data governance committee for data management plans.

Step 3. Research and innovation storyline

- Data for the research can be generated by **Researcher** self (see Storyline: Generating research data) and/or found in existing health data sources (see Storyline: Search data in metadata).
- **Researcher** sends a request for the use of the found data from existing health data sources, using **Research data request service**.
- **Research data request service** facilitates and administers process for data user to request selected dataset(s) from the **Central data provider review committee**.
- The **Central data provider review committee** verifies whether the desired analysis to be carried out meets the conditions of use of the requested dataset(s).

Step 4. Research and innovation storyline

- **Research data request service** makes data request to data provider(s), with
 - Data user identification
 - Mention of desired dataset(s)
 - Research question to be applied, with specific analysis
 - **Analysis processing environment** to be applied
 - Approval of **Local review committee launch research**
 - Approval of **Central data provider review committee**
- Data provider(s) check(s) the request against the authorisation metadata of **Data Preparator** terms of use.
- Data provider(s) check(s) the request on a veto of the **Local review committee data providing**

Step 5. Research and innovation storyline

- **Research data request service** links the result of the assessment back to the data user.
- **Researcher** combines the self-generated data and/or requested data from existing health data sources with a pseudonymization service in a **Central analysis processing environment**, or by retrieving results from federated analysis performed with the applicant.
- **Researcher** uses the data found and/or requested for the study.

Step 6. Research and innovation storyline

- **Researcher** selects the tool from the **Tool provider** to record research data and results at **Data producer** or **Data archiver** of the **Data holder** (the organization for which the researcher works).
- Selected tool is configured as a **Data producer** or **Data archiver** at a selected **Data holder**.
- **Local review committee data providing** of the relevant **Data holder** indicates under which conditions of use, data may be recorded at **Data producer** or **Data archiver** for multiple use.
- **Researcher** creates research results.

Step 7. Research and innovation storyline

- Based on the data-centric principle, **Data Preparator** separates the original health data from the original application (**Data producer** or **Data archiver**) and, where necessary, prepares the health data for multiple use (including for use for research and innovation) in a persistent data platform.
 - If the **Data producer** or **Data archiver** has not made the data sufficiently FAIR, the **Data Preparator** will, where possible, further provide it with (additional) (1) Encodings (using the correct coding and modelling according to of language determined by the **Data governance committee**) and context information, (2) Terms of use according to the applicable data policy of the **Data provider** (3) Review from the Terms of use service (4) Quality or usability characteristics (5) Provided with Version control where necessary
 - If the **Data producer** or **Data archiver** has made the data sufficiently FAIR, the **Data Preparator** will implement the data-centric principle from a persistent data platform, which acts as a transport layer towards the **Data provider**.

Step 8. Research and innovation storyline

- **Data provider** sets up the delivery process of the dataset based on the terms of use described in the authorisation metadata of the **Data Preparator** and the target transformation.
- **Data provider** registers and publishes the metadata of your choice
 - via a FAIR datapoint
 - with a data catalogue
- **Data provider** reports the FAIR data point used to the **Addressing service**.

Comments and questions storyline research and innovation

Question / Comment	Reaction
How does research generate data? Will this remain within Health-RI hub or will the research in question go to a "bronze system" of a research (healthcare) institution?	Researcher selects tools (EDC), which are configured as Data Producers for a selected Data Holder. Local review committee data providing of the Data Holder concerned indicates under which terms of use data as a Data Producer may be recorded for multiple use.

Research and innovation storyline



Questions?

Storyline: Making data available



- Roles
- Different steps from storyline
- Comments and questions received
- Your questions

Step 1. Storyline: Making data available

- **Data producer** asks the **Local review committee data providing** whether the data may be offered
- **Local review committee data providing** determines, based on the predefined data policy, whether and under which user conditions the data may be offered.
- **Data governance committee**
 - Indicates for previously defined data how to make it FAIR, which coding and modelling to apply
 - Determines for new data how to code and model it

Step 2. Storyline: Making data available

- **Data producer** produces, creates and stores the original health data on the basis of agreements made and guidelines for, among other things, storage of data, ownership, intellectual property, terms of use, selection and retention. For original data sources, where the user does not have the basis to know personally sensitive information (research data sources), the original health data is pseudonymized or anonymized. **Data producer** produces data preferably with
 - Context information
 - Terms of use drawn up by **Local review committee data providing** and combined with the Terms of use service
 - Coding and modelling according to unity of language determined by the **Data governance committee**
 - Quality and usability features
 - Version control

Step 3. Storyline: Making data available

- **Data Preparator** and **Data Producer** use an existing (reusable) interface or create new reusable interface (preferably in collaboration with software supplier data producer), between **Data producer** and **Data Preparator**.
- Based on the data-centric principle, **Data Preparator** separates the original health data from the original application and prepares the original health data, where necessary, for multiple use (including for use for research and innovation) in a persistent data platform.
 - If the **Data producer** has not made the data sufficiently FAIR, the **Data Preparator** will, where possible, further provide it with (additional) (1) Encodings (using the correct coding and modelling according to of language determined by the **Data governance committee**) and context information, (2) Terms of use according to the applicable data policy of data provider (3) Review from the Terms of use service (4) Quality or usability characteristics (5) Where necessary, provide version control
 - If the **Data producer** has made the data sufficiently FAIR, the **Data Preparator** will implement the data-centric principle from a persistent data platform, which acts as a transport layer towards the **Data provider**.

Step 4. Storyline: Making data available

- **Data provider** sets up the delivery process of the dataset based on the terms of use of the **Data Preparator** and the target transformation.
- **Data provider** registers and publishes the metadata of your choice
 - via a FAIR datapoint
 - a data catalogue
- **Data provider** reports the FAIR data point used to the **Addressing service**
- **Data guide** retrieves the latest state of affairs from affiliated data catalogues and/or FAIR data points on a regular and/or triggered by a search. (see Storyline: Search data in metadata).

Comments and questions received Storyline: Making data available (1/2)

Question / Comment	Reaction
So this process is just as heavy for an individual researcher who wants to share his dataset as it is to share data sets running out of healthcare systems?	Yes, in principle it is
Which level FAIR should be used?	In principle, that is up to the Data Holder. Data user can set requirements for level of FAIR to include data id catalogue or not.
Note: Step 3 (connection persistent data platform and data preparation) is in practice the biggest challenge to make the Health-RI a success. This step is within the responsibility of the healthcare institutions, I think it is important that sufficient attention is paid to this from Health-RI.	Agree

Storyline: Making data available



Questions?

Next steps



- Process feedback
- June 21 will finalize the texts with final responsibility.
- After that, the heads of the regional nodes of the Health-RI ecosystem give the final verdict on the components that concern the nodes.
- Around 12 October – Health-RI Conference – Open consultation V2.0