

RSO Zuid-Limburg data liberation strategy

Auteur: R.Nienhuis, J.Holslag, I. Schoonbrood

Datum:15-07-2026

Versie:2.0

Status: Voorlopig vastgesteld

Version history

#	Date	Remarks	Author
1.0	12-10-2025	Initial version	Remko Nienhuis, Igor Schoonbrood, Joost Holslag
1.1	22-01-2026	Added federation and modularity as explicit elements of RSO vision	Remko Nienhuis
1.2	08-07-2026	Textual improvements	Remko Nienhuis
2.0	15-07-2026	Approved by Management Team RSO Zuid-Limburg	Wim Eurlings

Table of Contents

Introduction	5
Mission	6
Vision.....	7
Vendor neutral storage and communication of health-data.....	7
Federation	7
Technical and commercial modularity: stay away from ‘total solutions’	8
Clear privacy design strategies	8
Rules of execution.....	10
Rule #1: Clear vision	10
Rule #2: Sustainability	10
Rule #3: Quality	10
Rule #4 Reuse	11
Rule #5: Being responsible	11
Target architecture	12
Data-ecosystem roadmap	14
Phase #1	14
Phase #2	16
Phase #3	17
Phase #4	18
Phase #5	19
Phase #6	20
Moving from application-centric to data-centric.....	21
Challenges of data-centrism	21
Current EHR-systems won’t adopt vendor neutral storage anytime soon	21
Healthcare domain- and use case specific modelling	21
Data is locked up in silos.....	22
Dutch healthcare providers lack data literacy	23
Roadmap to data-centrism	23
Moving from centralized to distributed	25
Challenges of federating disparate CDR’s	25
Creating and implementing a standard for openEHR federation	26
Record identification	26
User authentication	27
Patient consent.....	27
Resource distribution	28
Different openEHR implementations	29
Storing data in a distributed CDR	29
Where do apps reside?	29
The distribution roadmap	30
Silo integration.....	32
Data flows	33

Context-launch.....	34
FHIR out.....	36
FHIR in	37
Replicate 2 CDR.....	38
Replicate from CDR	41
<i>App Integration</i>	<i>43</i>
<i>Local integration strategies</i>	<i>44</i>
Responsibilities	45
Regional configuration management	45
Regional Data Governance	46
Use case integration design	46
<i>Vendor management.....</i>	<i>48</i>

Introduction

This document provides an overview of how the Zuid-Limburg data-ecosystem will evolve over time. It provides a brief description of the vision that underpins the data-ecosystem and the ‘principles of execution’ that guide its development. It provides a high-level roadmap and insights in how the ‘double movement’ to a distributed and data-centric platform will be implemented in a step-by-step fashion.

The document serves as an introduction to the Zuid-Limburg data-ecosystem and its evolution, for anyone that is involved with its development. It’s written in English for reasons of accessibility: many of the vendors involved with the development of the data-ecosystem are international vendors, as is the openEHR community that provides valuable insights and experiences. First and foremost, the document is targeted at employees of RSO Zuid-Limburg and its participants, to provide clarity and focus. It will serve as a basis and framework for developing use cases and required integrations with existing systems, the so-called local integration strategies.

Like the data-ecosystem will evolve over time, so will this document. Future updates will include information about Identity and Access Management (IAM) and about how the data-ecosystem will incorporate (future) developments like expected legislation (such as EHDS), new standards and national agreements (such as new NEN-norms), national facilities such as MITZ and Zorg-AB and the work being done by the CumuluZ initiative. Being a living document, it will reflect these developments in Health IT (HIT) and Health Information Exchange (HIE) in coming iterations.

Mission

The Zuid Limburg data-ecosystem aspires to provide patients and healthcare professionals with the information required to stay- or become healthy and to provide high quality healthcare. It supports the creation, collection, transfer and use of health- and healthcare data across providers.

Vision

Key elements of the RSO vision on the data-ecosystem are:

Vendor neutral storage and communication of health-data

Data should be stored and made available in a vendor neutral format as much as possible, and independent from the eventual utilization of that data. This will:

- Prevent Health-IT vendors from engaging in information blocking strategies
- Prevent vendor lock-in and create a sustainable ecosystem for innovation and transformation
- Lower costs associated with continuous extraction and transformation of data
- Ensure the availability of data for future and unforeseen use cases and systems
- Ensure data- and application portability as defined by the Apperta foundation¹. This will minimize data conversion efforts and prevent data loss when changing or switching EHR systems
- Creating a lifelong longitudinal health record for the Zuid-Limburg citizens

RSO utilizes both FHIR and openEHR standards as suggested Tsafnat et al² and recognized by Nictiz³

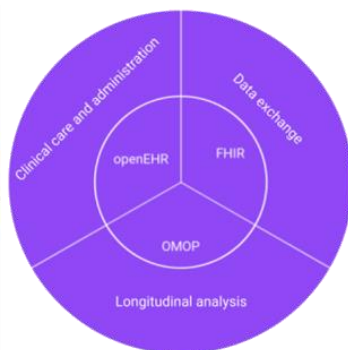


Figure 1: Different standards respond to different needs (Tsafnat et al)

Federation

Disparate sources of vendor neutral data will be ‘federated’ as a single queryable ‘virtual source of truth’. Federation requires unambiguous information modelling throughout the

¹ https://apperta.org/assets/Apperta_Defining_an_Open_Platform.pdf

² Tsafnat G, Dunscombe R, Gabriel D, Grieve G, Reich C. Converge or Collide? Making Sense of a Plethora of Open Data Standards in Health Care. J Med Internet Res. 2024

³ Verder met zibs. Voorkeursscenario nader toegelicht. Nictiz, 2024

data ecosystem as a whole and therefor solid data governance. RSO Zuid-Limburg will promote the development of international open specifications for openEHR federation within the openEHR technical community, that can be adopted by international openEHR-based vendors and opensource programs.

Technical and commercial modularity: stay away from ‘total solutions’

The data-ecosystem will be composed of technically and commercially independent modules and will avoid best-of-suite ‘total solutions’. Modularity will prevent high replacement costs associated with best-of-suite solutions and will promote post-procurement-competition. Modularity plays a crucial role in vendor- and risk management, preventing lock-in, unmanageable costs and innovation stagnation.

Clear privacy design strategies

All health data should be processed with clear privacy design strategies in mind. RSO bases the privacy design of the data-ecosystem on the 4 data-oriented design strategies formulated by Jaap-Henk Hoepman in the Little Blue Book:

- **Minimize:** Limit as much as possible the processing of personal data. This strategy will favor designs that facilitate fine-grained querying of data ‘on demand’ over exchanging predefined large (multi-purpose) datasets.
- **Separate:** Logically or physically separate the processing of health data collected by different healthcare providers as much as possible. This strategy will favor distributed design over centralized design.
- **Abstract:** Limit as much as possible the detail in which health data is processed. Again, this strategy will favor designs that facilitate fine-grained querying over exchanging predefined large (multi-purpose) datasets.
- **Hide:** Protect health-data, making sure it cannot be accessed or observed by parties without the appropriate authorization to do so. This strategy favors designs in which health-data is separated from personal identifying data as much as possible and requires enforcement of fine-grained access policies as well as obfuscation/ encryption of data in rest and in transit.

These key elements of the RSO vision match with the ‘double movement’ suggested by D&A Medical Group⁴ that can be depicted by Figure 2.

⁴ Onderzoek landelijk netwerk van infrastructuur voor gegevensuitwisseling in de zorg, D&A Medical Group, 2023

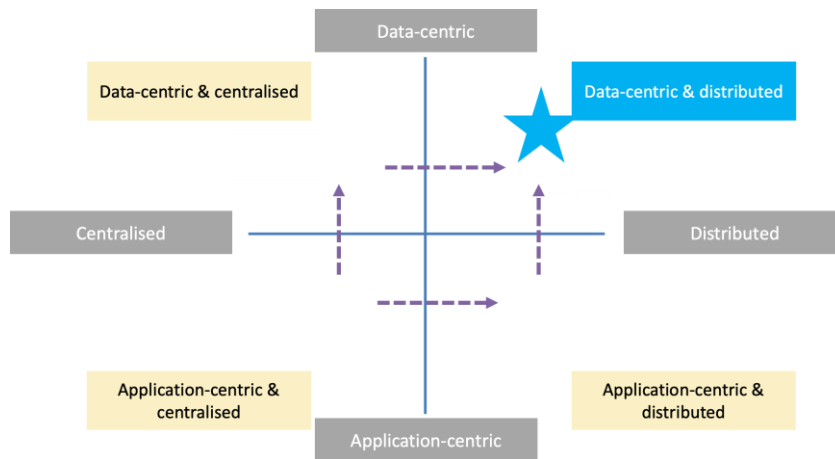


Figure 2: Double movement towards datacentric and distributed design

More information on the RSO vision can be found in [Position Paper Regionaal data-ecosysteem Zuid-Limburg, RSO Zuid-Limburg, 2024] and [IZA Transformatieplan RSO Zuid-Limburg, RSO Zuid-Limburg, 2025]. Other elements of the RSO design vision include:

- The use of national common infrastructure for services like patient consent (MITZ) and authentication (Dezi)
- The application of international standards by default, and national standards when needed based on a comply or explain regime
- The use of specific frameworks for trust-over-ip, such as NUTS and TWIIN.
- The use of specific profiles such as Smart App Launch and Smart backend services.
- Alignment with LDN (Dutch National Health Network) and LVS (Dutch National trust framework)

Rules of execution

RSO Zuid-Limburg, together with her technology partners, will create a regional health data-ecosystem and an associated ecosystem of apps on top of that. Five rules guide the execution of the program.

Rule #1: Clear vision

A clear vision on the design of the data-ecosystem and how it will evolve over time

Making sure that all RSO team members, participant board members and key users understand, embrace and promote the vision behind the design of the data-ecosystem and the reason behind choices and principles. This manifest aims to provide more insight into that design and how it will evolve over time.

Rule #2: Sustainability

Meeting today's demands while investing in the long-term vision.

Healthcare is facing challenges that require safe utilization and sharing of health data, rather sooner than later. Realizing the long-term vision on the data-ecosystem should not stand in the way of creating short-term value for healthcare professionals and their patients. On the other hand, short-term value should not prevent from achieving long-term goals such as preventing information blocking and vendor-lockin. This means that instead of waiting for the perfect solution, short-term design decisions might depart from the long-term design vision. Whenever they do, implications and risks for the long-term vision are identified and a long-term strategy is created that addresses these implications and risks.

This principle is in line with the aforementioned 'double movement' of Figure 2, but reaches further than the aspects of 'data-centrism and distribution'. Other key elements include other strategies and principles, such as the privacy design strategies, the use of specific standards such as NUTS and Smart App Launch and alignment with LVS and LDS.

Rule #3: Quality

Quality over features and timeline.

Creating the data-ecosystem will involve making decisions on the number of features to deliver within given timelines. Due to the sensitivity of working with people's health and health-data, as well as the importance of health-professional adoption, quality is non-negotiable. This means that we must pick between the number of features or the timeline of projects and never compromise quality.

Rule #4 Reuse

Utilize existing systems and tools over creating new infrastructure

Existing API gateways, messaging brokers and other HIE infrastructure is reused as much as possible. This includes participant specific infrastructure, as well as national infrastructure such as the LSP and national common services such as Mitz.

Rule #5: Being responsible

With influence comes responsibility

Responsibility and accountability should be placed on the shoulders of those that have the corresponding influence. Making data in existing systems available to the ecosystem is the responsibility of the individual RSO participants, simply because they alone have access to EHR systems and their vendors.

Target architecture

The (conceptual) target architecture of the RSO data-ecosystem can be described in three layers, each layer being composed from disparate technical or functional components that are integrated using (inter)national standards and are commercially independent:

- **Data availability layer**
Each healthcare provider that connects to the data-ecosystem is supposed to implement components that together facilitate vendor-neutral storage and access to data.
- **Platform layer**
This layer is implemented using technology and components that facilitate data federation and aggregation and the development of meaningful applications on top of that federated data. Either using commercial or open-source components of national services such as MITZ.
- **Process**
This layer is composed out of conceptual components that describe functional requirements that are derived from the IZA transformation plans in the region



Figure 3: Data-ecosystem Zuid-Limburg (conceptual) target architecture

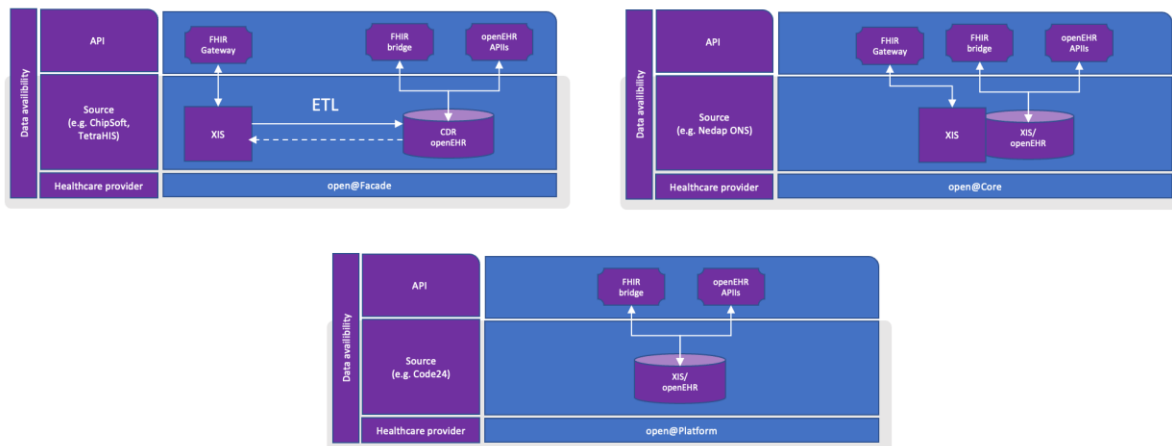
The data availability layer exposes data in openEHR and FHIR format. RSO distinguishes three different architectures:

- **Open@Facade**
Systems that store all data in a vendor specific format use this architecture. Data is replicated to an openEHR compliant CDR using a sync-agent, a regional ruleset and custom transformations. Data is exposed through AQL for fine-grained querying and FHIR for RESTfull coarse-grained access.
- **Open@core**
This strategy can be applied to systems that are able to store part of the data in

openEHR format. Depending on the required data, this architecture must be combined with the open@Facade architecture.

- Open@platform

This strategy can be applied to systems that store all data in openEHR format.



Data-ecosystem roadmap

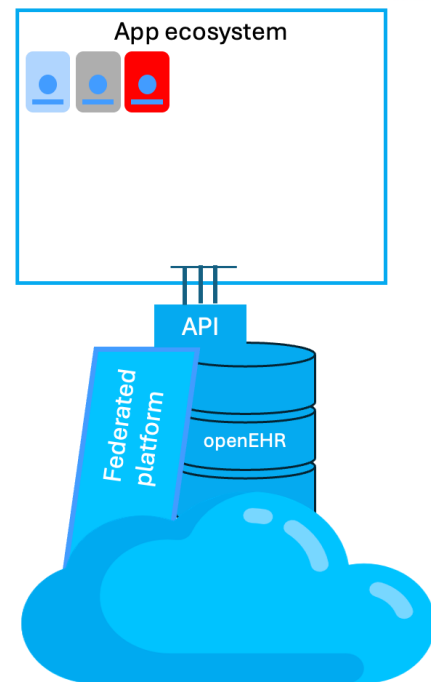
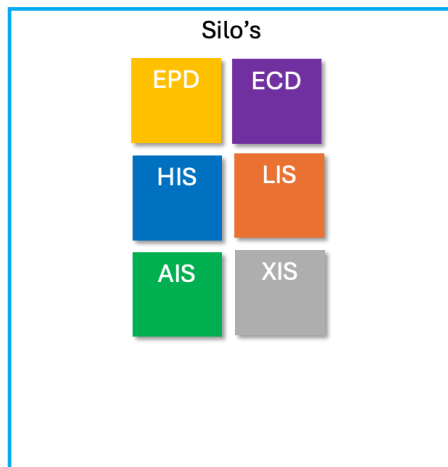
The data-ecosystem will evolve over time. A long-term high-level roadmap will guide the construction of the data-ecosystem. Consecutive chapters will provide more in-depth detail on the implementation of the phases described in the data-ecosystem roadmap.

Challenges that stem from the creation of a distributed system are described more comprehensively in chapter 'Moving from centralized to distributed' and those that stem from the datacentric approach in 'Moving from application-centric to data-centric'. Specific EHR- and App integration strategies will be the focus of chapters 'EHR integration' and 'App Integration'. Connecting RSO participants to the data-ecosystem will be the focus of chapter 'Local integration strategies'.

Phase #1

During phase #1:

- A centralized openEHR CDR platform is in place
- A small number of apps is in place, supporting regional use cases
- Some (if not all) of these apps are built by citizen developers using no-code/ low-code tools.
- Apps store and retrieve data from the CDR platform
- There is no integration with existing EHR's at all
- RSO starts with executing the data-centrism roadmap (see Roadmap to data-centrism).



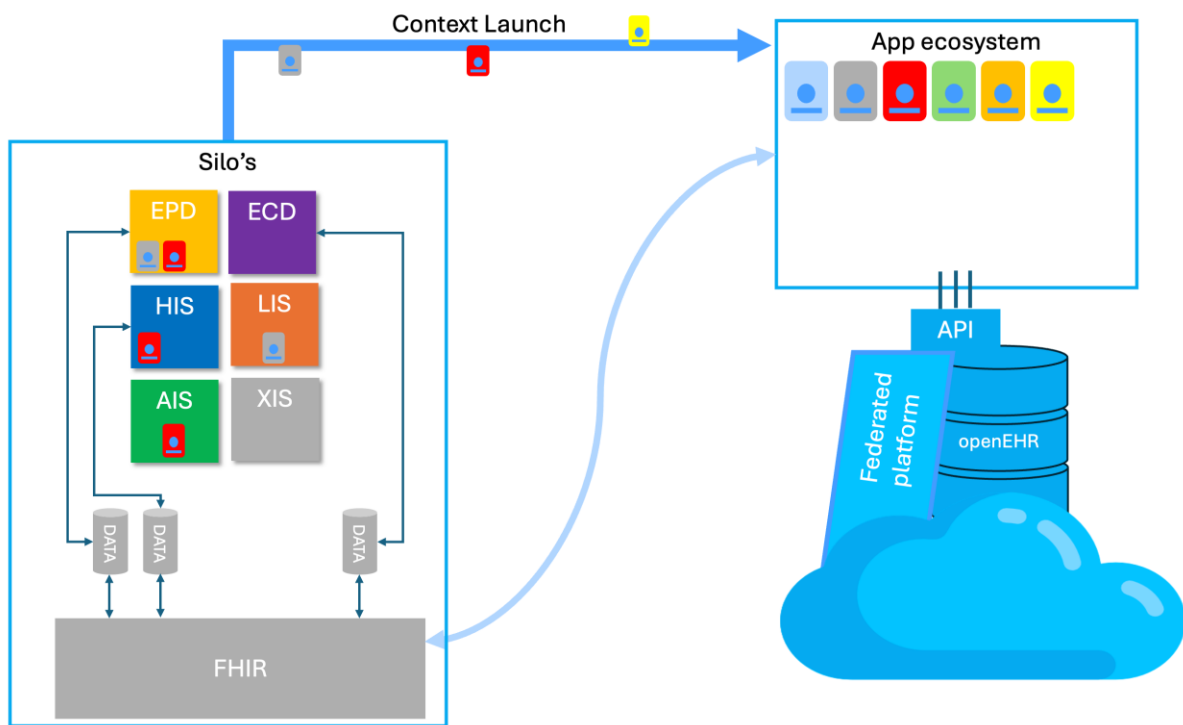
Challenges include:

- User adoption with minimal integrations might prove challenging

Phase #2

During phase #2:

- Simple FHIR based communications are set-up between EHR systems and apps, essentially providing patient demographic data to use with context launch
- Apps are visually integrated within EHR systems using Smart app launch



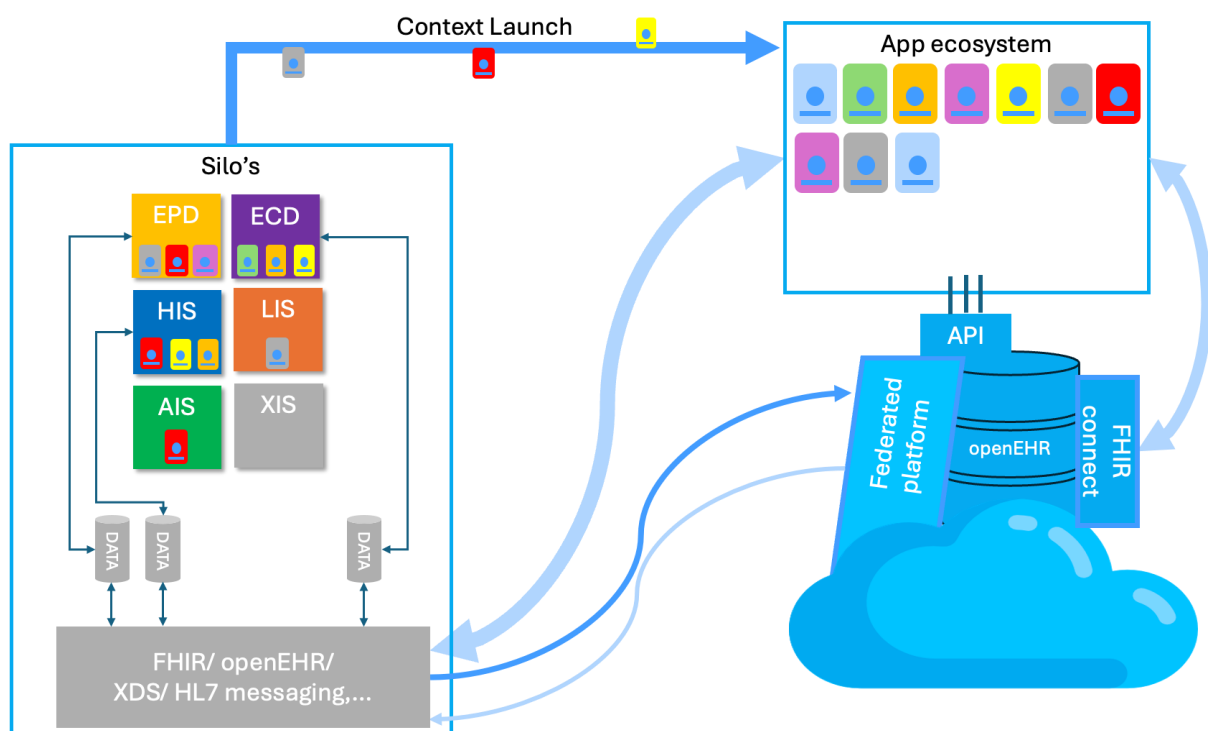
Challenges include:

- Most EHRs don't implement standards for context launch such as Smart app Launch
- Even creating FHIR connectors for only demographic data might prove a challenge for some EHRs

Phase #3

During phase #3:

- The number of apps/ use cases grows
- As is the number of Context Launch integrations
- Most existing apps on the Dutch market don't support openEHR. OpenFHIR and the underlying FHIRConnect standard will help transform FHIR to openEHR and vice versa
- The platform is federated, meaning that logically or physically separated CDR instances can be queried as one
- Apps start to use more diverse FHIR data coming from provider EHR systems. A regional FHIR federation platform will cumulate data from different sources
- Some apps use a FHIR connector to the CDR to retrieve and store data
- Data is replicated from EHR systems to the CDR platform. This is how we gradually make siloed data available in a vendor neutral (openEHR) format
- When necessary for the correct functioning of EHR systems, for example to show notifications about treatment limits, data is replicated from the CDR to EHR systems



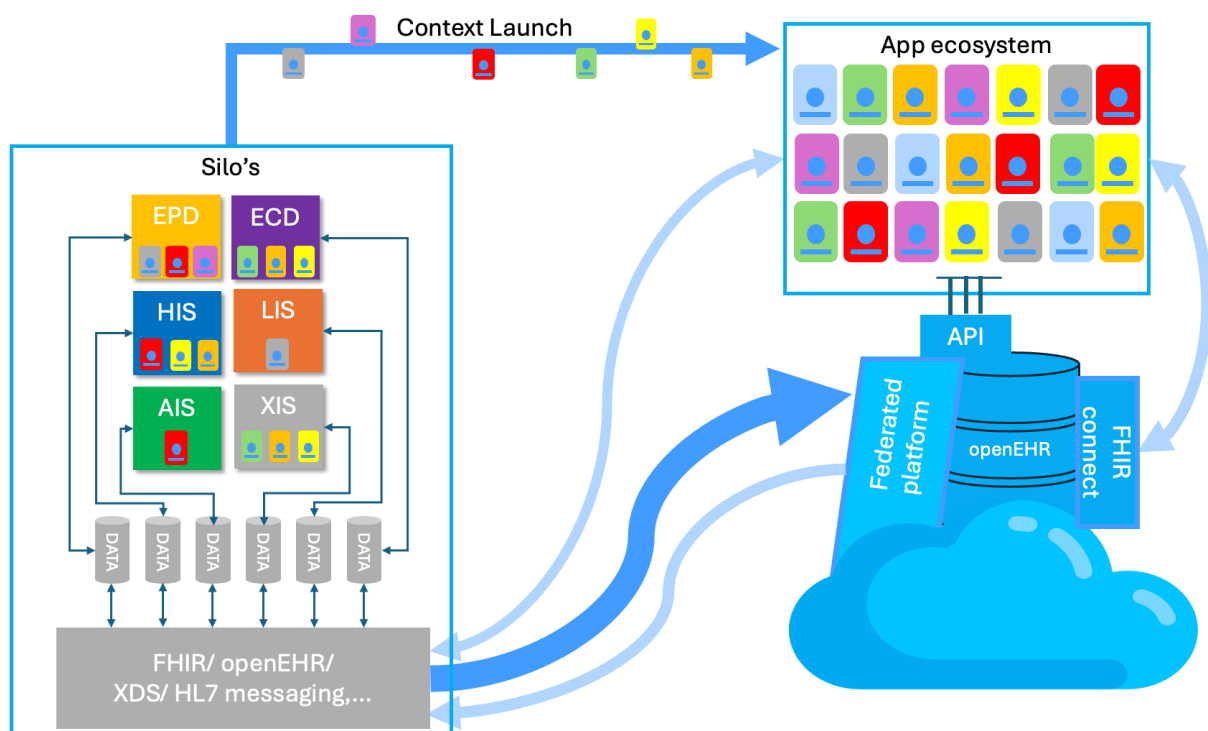
Challenges include:

- Creating a replication infrastructure will prove challenging for some participants, even with the help of RSO and technology partners
- FHIR connectors to EHRs remain challenging
- Implementing federation for both openEHR and FHIR
- Preventing 'broadcasting' personal data (BSN) to participants that have no treatment relationship with a given patient (localization)
- Assuring a legal base for data exchange, such as patient consent
- Moving from a centralized approach to a distributed approach will include challenges like distributing existing data and creating policies that honor the dynamics of distributed information management, such as who can create, update or delete data in a distributed environment

Phase #4

During phase #4:

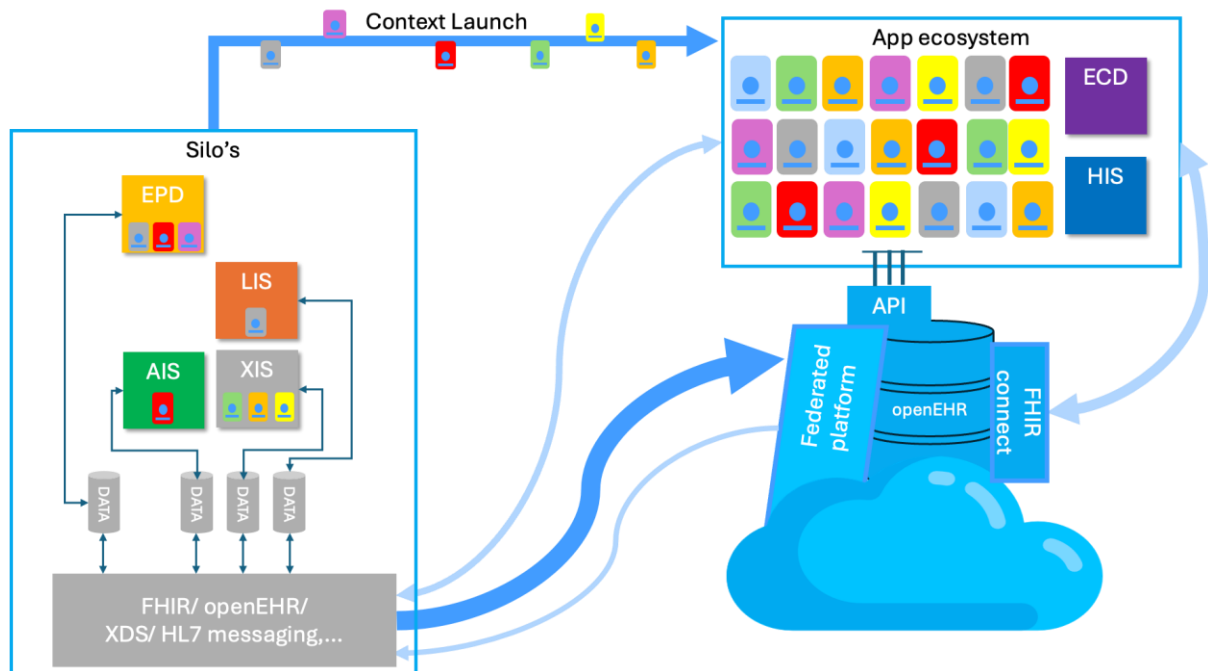
- The number of apps/ use cases grows even more
- As is the number of Context Launch integrations
- More data is replicated from EHR systems to the CDR platform
- A little more data is replicated from the CDR to EHR systems
- Apps created with the platforms no-code/ low-code tools are shared between (inter)national programs using the platforms marketplace



Phase #5

During phase #5:

- Some EHR systems will start to use openEHR as a storage format and store data directly in the federated CDR.
- This will decrease the number of FHIR integrations with apps and the replication of CDR data to EHR systems.



Challenges include:

- Getting XIS (silo) vendors to implement vendor neutral storage (such as openEHR) will be difficult and might even never happen⁵

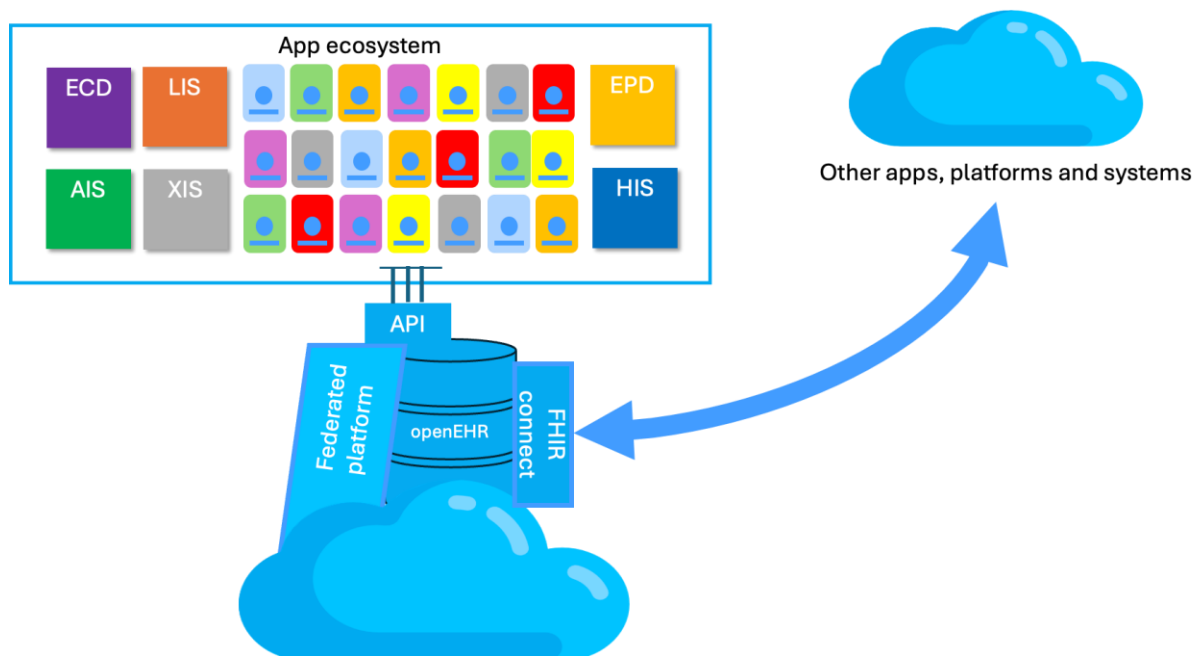
⁵ Some Dutch vendors are going through this transition though (Code24 and Nedap). Internationally the Tieto Evry EHR, that is widely in use in the Nordics, is based on openEHR

Phase #6

During phase 6:

- All EHR systems and apps will start to use openEHR as a storage format and store data directly in the federated CDR.
- FHIR is used as a data exchange mechanism with apps that don't support openEHR and platforms that are based on different designs
- FHIR is used to connect the data-ecosystem to the national health network (LDN/ CumuluZ)

Note: The use of FHIR to connect the national Health Network (LDN), could (should!) be realized in an earlier stage, depending on when the LDN is available.



Challenges include:

- A scenario where all XIS (silo) vendors use vendor neutral storage will happen only if this is enforced by regulations.
- Using FHIR to connect to other regions depends on the existence of technical agreements, information standards and services that together form the Dutch LDN. At the time of writing, it's still hard to tell which directions LDN will take.

Moving from application-centric to data-centric

An important part of the vision behind the data-ecosystem, is that data should be stored and made available in a vendor neutral format as much as possible, and independent from the eventual utilization of that data. Separating data from its use in 'applications' is what we call data-centrism.

Challenges of data-centrism

Current EHR-systems won't adopt vendor neutral storage anytime soon

Some (inter)national vendors moved from using legacy information models to openEHR or FHIR based storage of clinical information, but they represent a small minority. Fortunately, Nedap, a large Dutch vendor that operates in different healthcare domains, is making the switch to vendor neutral storage of (clinical) information. A smaller Dutch vendor, Code24, provides an EHR that was based on openEHR from the start. In the Nordic countries Cambio, TietoEvry and Dips have based their new generations of EHR systems on openEHR.

Still, without any further legislation⁶, most EHR vendors won't move to vendor neutral storage, due the high costs of such a transition, the competitive advantage they get from utilizing information blocking strategies and the lack of customer demand. This means that extracting data from- and updating data in- these 'siloes EHR-systems', will remain a costly enterprise. It involves the endless creation of (and paying for) use case specific extractions and transformations, and the utilization of a variety of custom APIs.

Replicating (extracting and transforming) data from silo systems to a vendor neutral storage platform such as an openEHR CDR, will remain an important strategy for anyone who wishes to reuse clinical data at a large scale⁷. Utilizing such a strategy means that extracting and transforming data is only done once, after which data can be reused for many different use case, including secondary use.

Healthcare domain- and use case specific modelling

Although in the Netherlands a long-standing tradition of clinical modelling exists (the so called zorginformatiebouwstenen), clinical models are often (partly) use case- and

⁶ Note that the EU data strategy and Data act push for vendor neutral storage of data by cloud providers

⁷ Note that this is also one of the strategies utilized by the CumuluZ initiative and an important strategy for FAIRifying clinical data.

healthcare domain- specific. Next to that, they are often country specific and differ from international models such as the international openEHR archetypes.

The Dutch healthcare system is highly fragmented, each domain using customized information models, coding systems and information recording practices, supported by a large number of different EHR systems and smaller vendors. This, off course, explains at least some of the divergence in clinical modelling for the Dutch situation. Fortunately the need for harmonizing models and coding systems accross healthcare domains is becoming more and more recognized by policymakers and healthcare professionals. Still policies and tools to help with the harmonization are missing.

The Dutch zib initiative lacks modelling formalisms such as a formal (machine enforceable) relationship between clinical concepts and their use in clinical practice (such as the relationship between archetypes and templates in openEHR), which decreases zib reusability across use cases and domains. The lack of a stable common information model as a basis for zib-modelling, leads to divergent modelling and difficulties implementing the models in software. Last but not least zib-modelling lacks (enforcement of) clear (computable) rules for localizing, specializing and extending models and for rigid version control. This partly explains the differences with international archetypes but also leads to avoidable breaking changes between versions and use cases.

The ‘zib transformation’ community addresses some of these problems and released an updated vision on zib-modelling, informally called zibs 2.0. The zib 2.0 approach includes a vision on the relationship between openEHR archetypes, FHIR profiles and the intermediate (technology agnostic) zib models. Harmonizing openEHR archetypes and FHIR profiles is a key element of the zib 2.0 vision, but the vision still lacks a clear roadmap or methodology for doing so and addresses this as necessary future work.

A lack of use case- and healthcare domain independent models impedes a truly data-centric approach to interoperability.

Data is locked up in silos

Due to the siloed nature of most EHR-systems, exporting data from these EHRs and transforming it to a vendor neutral model can prove challenging or even impossible. Also, data inside these siloes is often not semantically compatible with (inter)national

clinical models. It has a different structure, uses different data types or different coding systems. Even when it's possible to extract this data, it might be impossible to transform it to a common standardized format. This makes a data replication strategy challenging. More on synchronizing data from EHRs with an openEHR CDR is covered in 'Replicate 2 CDR'.

Fortunately, the Dutch zib initiative did have some effect in harmonizing data registration in EHRs with standardized clinical models. The zib 2.0 efforts will hopefully build on that foundation and further improve the quality of data in siloes.

Dutch healthcare providers lack data literacy

Data literacy is the ability to create, read, understand, work with, analyze and communicate with data. Many Dutch healthcare providers lack an understanding of the importance of data and data-literacy, and give limited priority to improving data-literacy. This is partly due to the fact they have little influence on how data is handled by silo systems.

As a consequence, data-governance practices are confined, there is little sharing of knowledge and practices and articulation of customer needs is fragmented and unfocussed.

Roadmap to data-centrism

The roadmap to data-centrism will be built around five tracks that will be executed simultaneously:

- Harmonize data in silos with (inter)national clinical models
This 'data governance' track will build on the current Nictiz zib 2.0 strategy. It includes pushing standardization efforts in the Netherlands towards harmonizing zibs and other clinical models with their international counterparts, especially the international Clinical Knowledge Manager. The track also includes activities that will harmonize data models in silo-systems with these (inter)national clinical models, such as reconfiguring systems, adjusting registration practices and vendor management activities (see Vendor management).

An example of (re)configuring systems to harmonize data with (inter)national clinical models is the configuration of 'Questionnaires'. These can be created in

such a way that they are semantically interoperable with (inter)national archetypes.

- Open-up silos for replication strategies

Real-time access to data in silo-systems, by using (FHIR) APIs might prove difficult at times or insufficient for supporting regional use cases. One possible solution can be to replicate data and store it in the CDR associated with the healthcare provider that is responsible for the silo-system(s). See 'Replicate 2 CDR'

Support for FHIR questionnaires will prove crucial for opening-up silos. Many silo-systems store a significant amount of clinical data in the form of questionnaires. Making these available through FHIR APIs will mean a huge step forward.

- Improve data-literacy

Another 'data governance' track that focusses on creating awareness among healthcare professionals and senior staff on the importance of data and data literacy. RSO will also help participants with creating the ability to create, read, understand, work with, analyze and communicate with data.

- Push for vendor neutral storage

This track will include creating awareness about the importance of vendor neutral storage amongst policy makers and healthcare senior staff, but also vendor management activities such as customer need articulation and publicly praising vendors that lead the way.

- Focus on meeting today's demand

Principle of execution #2 states that RSO has to focus on today's demands, even when this leads to a deviation from the architectural vision (distributed and data centric).

Moving from centralized to distributed

Following the privacy design strategy ‘Separate’, the openEHR CDR platform will be implemented as a distributed platform. Data that is created by a specific healthcare provider will remain under the influence and ‘control’ of that provider as much as possible. This is in line with government policies regarding ‘centralized storage of data originating from disparate health providers at one storage location’⁸.

This means that:

- Data created by a healthcare provider will be stored in a CDR that is under the control of that provider alone (control in the sense of a GDPR ‘controller’)
- Disparate CDR’s will be separated either logically or physically
- Copying of data to other CDR’s, stores or EHR systems will be prevented as much as possible.
- It will be possible to query disparate CDR’s as if it were a single CDR. This requires these disparate CDR’s to be ‘federated’.

The openEHR specification provides an implementation of the privacy design strategy ‘Hide’ by separating personal identification data and demographic data from clinical data. In a distributed environment such as the Zuid-Limburg data-ecosystem, the demographic data will be implemented in a distributed fashion as well. So, there will be no centralized Master Patient Index (MPI) for the region, but disparate healthcare providers will control their own instance of such an MPI, which will be fed from the main EHR- or patient administration system used by that provider.

The implementation of a distributed (openEHR based) EHR-platform, especially for primary use of healthcare data, will introduce challenges. Although proof of concept technology is available, there are no standards available governing the subject of federation. Unsurprisingly, there is no international experience with large scale distributed EHR platforms. Problems that are previously unaccounted for may arise, leading to poor user-adoption and patient safety concerns. Therefore, the Zuid-Limburg data-ecosystem will carefully evolve from a more centralized design to a distributed design.

Challenges of federating disparate CDR’s

⁸ <https://www.rijksoverheid.nl/documenten/kamerstukken/2024/01/22/kamerbrief-over-landelijk-dekkend-netwerk-van-infrastructuren>

Creating and implementing a standard for openEHR federation

At the time of writing the openEHR standard does not provide a specification for federating disparate CDR's⁹. Creating such a specification will take time. Also, the openEHR community must decide on a deployment architecture for the federation design:

- Centralized design
This means the introduction of a centralized 'federation gateway' that distributes queries to disparate CDR's and collects and accumulates results. The main advantage of a centralized design is that it's less complicated than it's distributed sibling and less dependent on participant specific infrastructure.
- Distributed design
This means that each CDR (or an AQL Query facade in front of the CDR) can act as a 'federation gateway', and that each of these gateways is under the control of the healthcare provider that controls the CDR.

The distributed design is more in accordance with the privacy design strategy 'Seperate' but is more complicated than a centralized design. For example, each disparate 'federation node' requires knowledge about the other nodes (addressing) which requires a protocol to safely distribute node-information across the network (note that the Dutch NUTS community provides the means to do this).

The international openEHR community, is currently in the process of developing the required specifications. RSO Zuid-Limburg and its technology partners participate in the development.

Record identification

In openEHR a patient record is uniquely identified by a meaningless id. The EHR does not contain patient identifying attributes such as the patients BSN. In a distributed environment, a particular patient will have records in disparate CDR's that are identified by disparate record-ID's that essentially point to the same individual. This means that a single distributed query will introduce overhead, because a given BSN must be mapped to different record-ids by the API/ federation gateway(s) connected to each disparate CDR.

RSO and its technology partners will adopt a solution based on a distributed Master Patient Index, where each participant provides it's own MPI that holds a mapping from

⁹ Note that such a standard is not available for FHIR either

the BSN to the openEHR EHR_ID. This strategy has the following advantages:

- This strategy doesn't require a centralized MPI
- This strategy doesn't require a querying party to have any knowledge about the disparate EHR-ids.
- This strategy doesn't require extensions to the AQL language to take handle BSN to EHR-id mapping.

For use cases that are developed on top of the centralized CDR however, a mapping to the central EHR_ID will be pushed to the local MPI of the creator of the record as well as to a central PMIR.

User authentication

The distributed/ federated openEHR CDR will be used by/ queried by different healthcare professionals from different healthcare providers. Central storage of user information will be minimized to administrative personnel only. This means creating a 'trust' between disparate healthcare providers:

- Creating a PKI or using an existing PKI
- Agree on attributes that will be used to identify users and make authorization decisions
- Adjusting existing SSO protocols
- Create a Dutch implementation guide for the Smart App Launch profile that builds on the PKI and agreed identifying attributes
- Align with Dezi (the new national authentication framework for healthcare professionals)

Patient consent

A healthcare provider must only provide clinical data in response to a federated query when sufficient patient consent is available. In a distributed environment, each individual node is responsible for determining whether the required consent is available. EHDS will eventually change the landscape, by enforcing the 'opt-out' construct instead of the Dutch explicit consent construct. Still, each node in a federated network is responsible for checking possible opt-outs.

The preferred implementation of checking for opt-out would be to request a national service. Especially when scaling up the data-ecosystem and 'federating' it with other regional platforms, a centralized consent service will be preferable to disparate (partly overlapping) regional services. The Dutch Mitz service could provide the means to do so in the near future. Each RSO participant will be responsible for checking consent or opt-out before providing health information of a given patient.

Resource distribution

In a distributed openEHR CDR, each individual CDR instance requires access to the same openEHR archetypes and templates. Federation requires all disparate CDR instances to work with the exact same set of archetypes and templates, because a difference in data-schemas will break distributed queries.

Rigid procedures for manually harmonizing archetypes and templates across all federated CDR instances will suffice on a small scale. When the number of federated CDR's grow and the number of additions and updates to archetypes and templates increases, the need for automation arises. This can be implemented by:

- Creating a single CDR that contains no clinical data but acts as a 'master repository' for archetypes and templates.
- Blocking the API's for individual CDR's that provide administrative access to archetypes and templates.
- Each time an archetype or template is added or updated to the master-CDR, the master-CDR will distribute changes through a streaming system such as Kafka.

Physical versus logical separation

Disparate CDR's can be separated logically, utilizing the same storage and (R)DBMS but creating disparate patient-records for each logical instance, or they can be separated physically. In the latter case, each healthcare provider stores clinical and demographic data in isolated storage and (R)DBMS systems, possibly in disparate storage locations and data centers.

The costs of physical separation will be higher than logical separation. Theoretically, physical separation could provide a higher level of security because intruders will have to obtain access to different storage platforms and locations, possibly offered by different service providers. In practice, the added complexity and costs can also introduce security risks and create new attack surfaces.

This means that:

- RSO supports both logical and physical separation of federated CDR's, and leave the decision to individual participants
- The platform design must facilitate moving data from a logical separated CDR to a physical federated CDR, and the other way around. In general, bulk export and import functionality will provide the means to do so.

Different openEHR implementations

The Zuid-Limburg data-ecosystem might include openEHR based systems from different vendors with different capabilities. For example, the Dutch EHR vendor 'Nedap' provides an EHR system that partially stores data in openEHR format and supports a subset of the openEHR specifications and API's. Some hospitals in the region use an opensource CDR (EhrBase). These different implementations come with different capabilities (e.g. Nedap does not yet support AQL) and add their own variants of non-standardized features such as bulk-export, event streaming and low-code support. This means that the evolution/ distribution of the data-ecosystem will involve modifying existing systems and building custom solutions to overcome differences in (non-standardized) capabilities.

Storing data in a distributed CDR

Some distributed systems, such as blockchain, will distribute added or updated information to each node in the network. Smart protocols such as 'proof of work' are in place to achieve consensus on the correct state of the data, including the relationship with previous versions of that data, even when nodes are faulty, compromised or temporarily unavailable. In the distributed EHR, clinical data will not be distributed across the network (copying data is not compliant with the 'Seperate' design strategy), but new and updated data will only reside in a single CDR: the CDR associated with the healthcare provider that is the controller for that data. This means that there is no way to make sure that data in the network is absolutely consistent. For example, there is no way to be absolutely sure that an update on an ACP plan is actually an update to the latest version of that plan (stored inside another CDR, with another id and its own versioning).

Users will only add or update data to or in the CDR that is controlled by the healthcare provider that they currently work for. When querying for data distributed across disparate CDR's, the resulting dataset can be grouped and ordered in such a way that data coming from different healthcare providers can be interpreted and compared by the user. Due to the (international) lack of experience with large-scale distributed EHR-systems, implementation of such a design must be done with concern and in a step-by-step fashion. Problems that are unaccounted for can introduce user-adoption challenges or worse.

Where do apps reside?

Distributing the data across disparate CDR's does not automatically mean distributing the applications that use that data. Different deployment architectures are conceivable:

- Host apps on a central location/ data center. This means that although data is stored locally, it is still processed by a central party (data processor) and therefore not in accordance with the 'Separate' privacy by design strategy.
- Decentralize apps in such a way that each healthcare provider hosts its own set of apps and apps remain under control of the individual healthcare provider, independent from a central data-processor. Hosting multiple versions of the same app will increase cost and attack-surface.

RSO must decide on a preferred deployment architecture and roadmap for each connected app. Implications of the preferred design must be identified in collaboration with technology partners. Using apps that are hosted on a central location on distributed data, might require changes to those apps so they 'know' where to store data based on what healthcare provider the user is (currently) working for.

The distribution roadmap

The RSO design-vision embraces the 'Seperate' design strategy and therefore imposes distributing data and apps across a regional network of collaborating healthcare providers. Then again, 'principle of execution #2' emphasizes the need for 'Meeting today's demands while investing in the long-term vision'. Given the challenges that come with the construction of a distributed EHR-platform, the platform will slowly but surely evolve from a more or less centralized design to a completely distributed design.

RSO will, together with its technology partners, create a high-level roadmap to facilitate the movement from centralized to distributed. The roadmap will make a distinction between use cases that involve joint record creation (different healthcare providers working together on the same data, such as ACP data) and use cases that only involve viewing data created and managed by individual healthcare providers (such as results from examinations and tests):

1. Data that is created and managed by regional joint processes, such as ACP data, may be stored in a centralized CDR (which requires joint data controllership). If this is the case, a roadmap must be created that addresses the steps needed to distribute data and apps.
2. Data that is created and managed by a single healthcare provider, such as laboratory results or examination data, is only stored in a CDR controlled by that healthcare provider. RSO will not allow replication of such data to a centralized repository.

Next to the use case specific design decisions, the roadmap will address fundamental problems, such as:

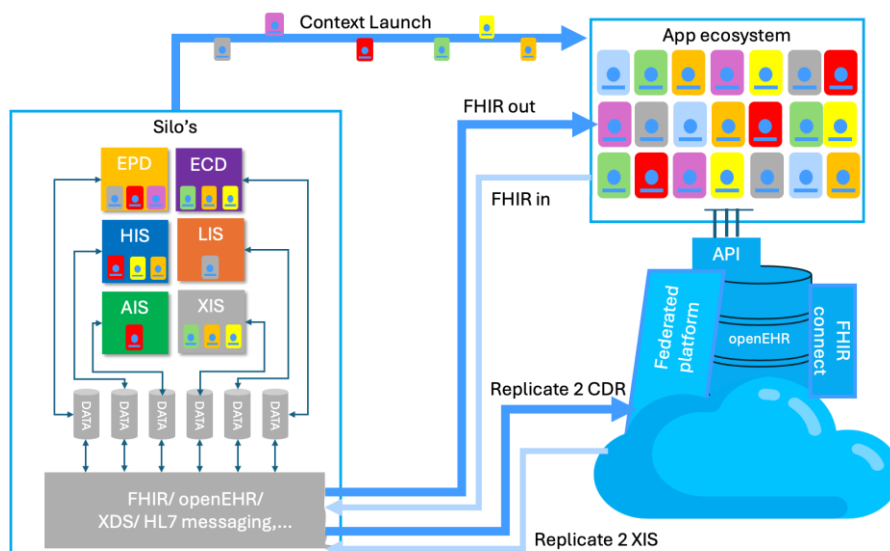
- The creation of a specification for federation that is supported by the openEHR community
- Services for patient consent and user authentication that are harmonized with national developments such as Mitz and Dezi
- Infrastructure for resource distribution
- Deciding on a strategy for logical or physical separation of data and apps
- Overcoming differences in CDR implementations
- Providing distributed solutions for common requirements such as addressing, localization and identification/ authentication.

Silo integration

Health- and healthcare data is collected by healthcare professionals and stored in proprietary systems and databases that are controlled by individual healthcare providers. Many of these systems behave as silo's: much of the data stored inside never leaves the silo and cannot be reused by other systems. The other way around is true as well; much of the data stored inside the silo can't be updated or created by other applications.

The RSO data integration strategy will be adapted for each disparate healthcare provider and tuned to the capabilities of their systems. In general, the strategy includes five dataflows:

- Context Launch
Context Launch allows apps to launch from inside or outside the user interface of an EHR system, providing the app with contextual information about the user and the selected patient.
- FHIR out
Providing FHIR data from silos to the data-ecosystem
- FHIR in
Apps use FHIR to store data in silo's
- Replicate 2 CDR
Using a plethora of preferably standardized mechanisms for (real-time) replication of data to the provider specific CDR instance
- Replicate from CDR
Using a plethora of preferably standardized mechanisms for (real-time) replication of data to the silo's



Next to implementing these 5 dataflows, the strategy will involve creating policies and mechanisms to control and improve data quality (data governance) and will involve vendor management focused on 'opening up silos' and encouraging vendor neutral storage.

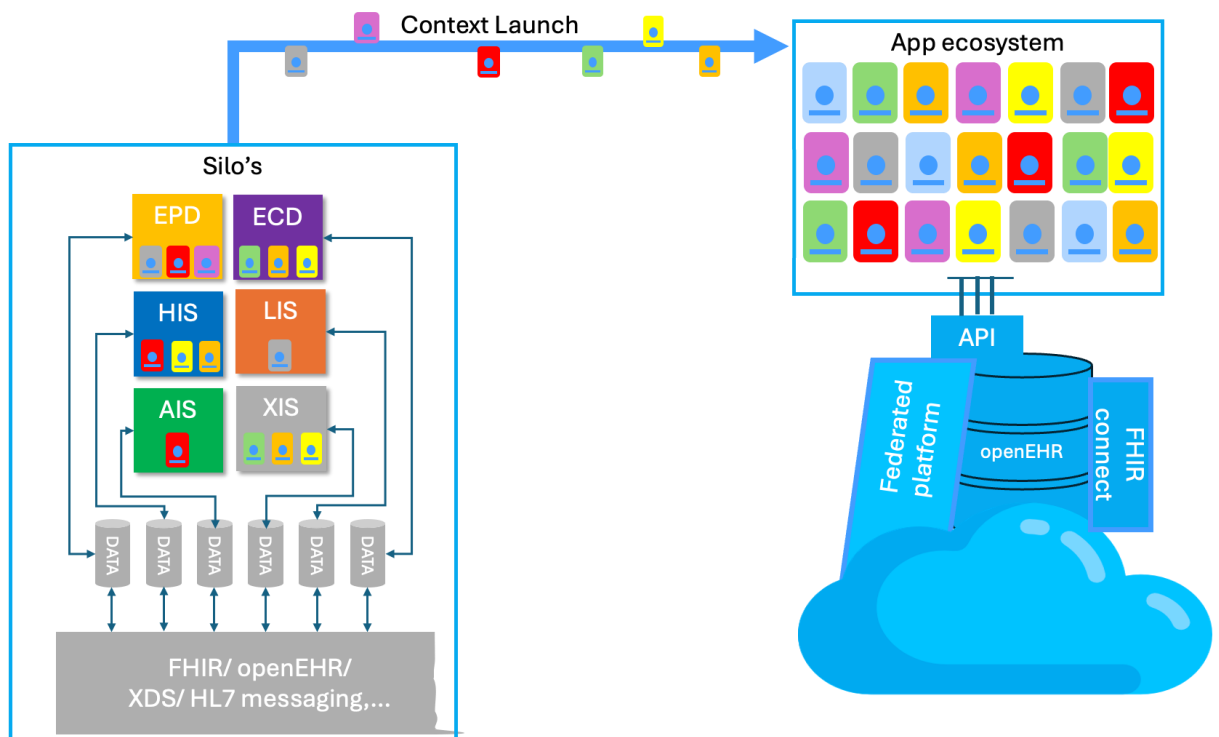
Data flows

Data flows are implemented by RSO participants, with help from RSO and technology partners and reusing existing infrastructure (HIE or EAI platforms) as much as possible.

Context-launch

This flow will involve:

- Use standards to visually integrate apps into the EHR systems used by healthcare professionals and share context (logged-in user and active patient)
- Implement silo-specific context launch solutions and transformations to standard profiles such as Smart App Launch



Challenges implementing this flow are:

- Most silos don't support international standards for context launch such as Smart App Launch
- The only standard that is supported in the Netherlands is 'Koppeltaal', which is not an international standard and therefore not widely supported
- The importance of supporting context launch is not widely recognized in the Netherlands. It might prove challenging to create an implementation guide for profiles such as Smart app launch that are consistent with national programs like Dezi.

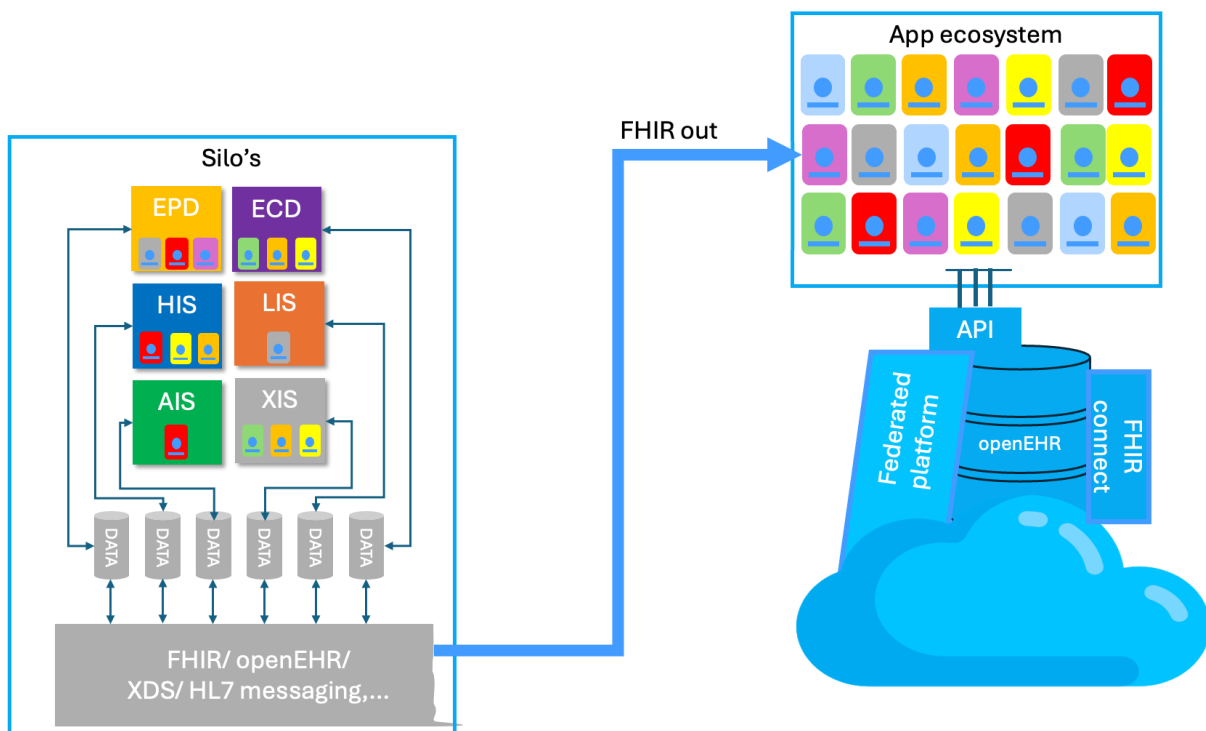
Depending on the capabilities of silo systems, the implementation of context launch will be less or more 'usable'. RSO Zuid-Limburg recognizes 4 levels of context launch, ranging from non-existing (stand-alone) to fully implemented (Maximize). For each integration the context launch level that matches the silo systems integration capabilities will be selected.

	● Low usability ● Moderate usability ● High usability STAND ALONE no integrations	EXPAND Single Sign On (SSO)	OPTIMIZE - Single Sign On (SSO) - Patient context sync - Access to zibs	MAXIMIZE - Single Sign On (SSO) - Patient context sync - Access to zibs and documents
ACCESS	● Login via 2FA/ DEZI	● No need to log in	● No need to login	● No need to login
PATIENT DATA	● Manual search/ entry	● Manual search/ entry	● Patient data synchronized	● Patient data synchronized
ZIBS (FHIR)	● No access to zibs in EMR	● No access to zibs in EMR	● ZIBs are linked to the application from EMR	● ZIBs are linked to the application from EMR
DOCUMENTS & IMAGES (FHIR/XDS)	● Manual upload /download	● Manual upload /download	● Manual upload /download	● Documents are linked to the application from EMR or XDS

FHIR out

This flow will involve:

- Exposing data from EHR-systems in FHIR format through APIs
- Reusing existing infrastructure as much as possible (e.g. FHIR APIs created for Dutch VIPP programs)



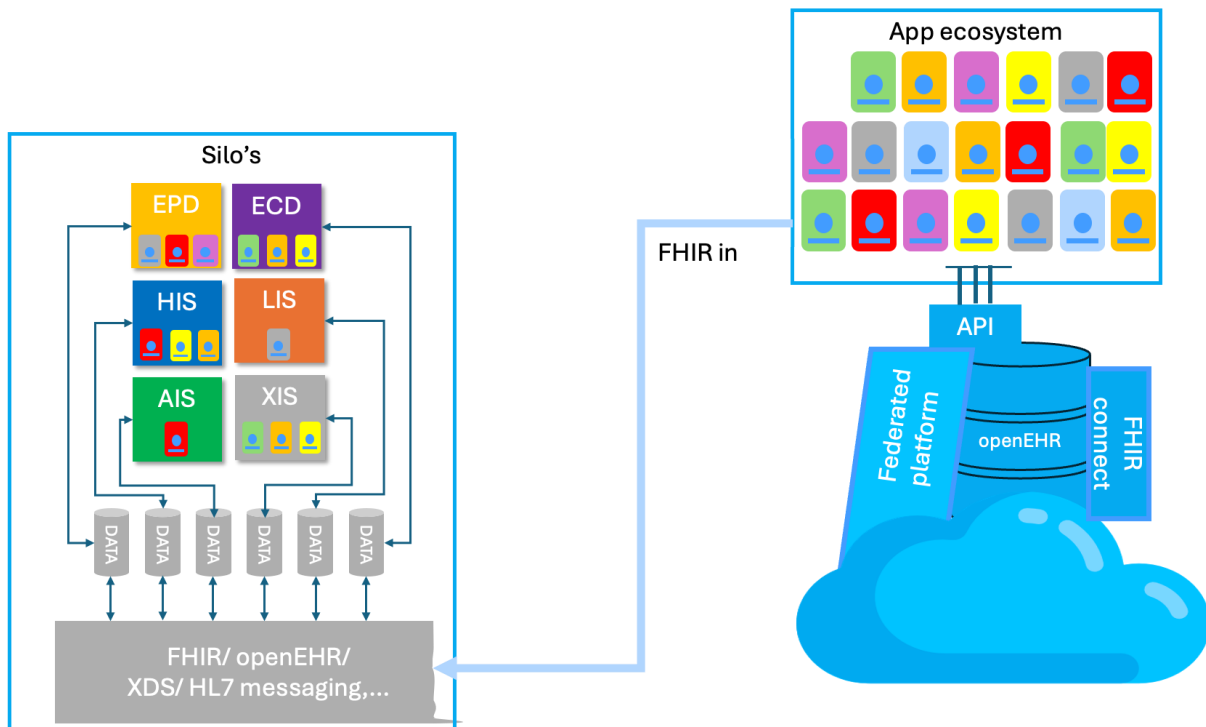
Challenges implementing this flow are:

- Data is inconsistent across healthcare domains and therefore incomparable. For example, coding systems used are incomparable and similar concepts are modelled differently
- Differences in API security – and trust models used across healthcare domains and VIPP programs limits their (re)usability
- Providing consolidated audit trail and (technical) logs

FHIR in

This flow will involve:

- Apps adding or updating data in silo's using FHIR posts or updates. The privacy strategy 'Minimize' plays a crucial role, not only for privacy reasons but also for practical reasons: adding or updating data in silos comes with significant costs.
- Reusing existing infrastructure, including communication platforms and older HL7 messaging standards



Challenges implementing this flow are:

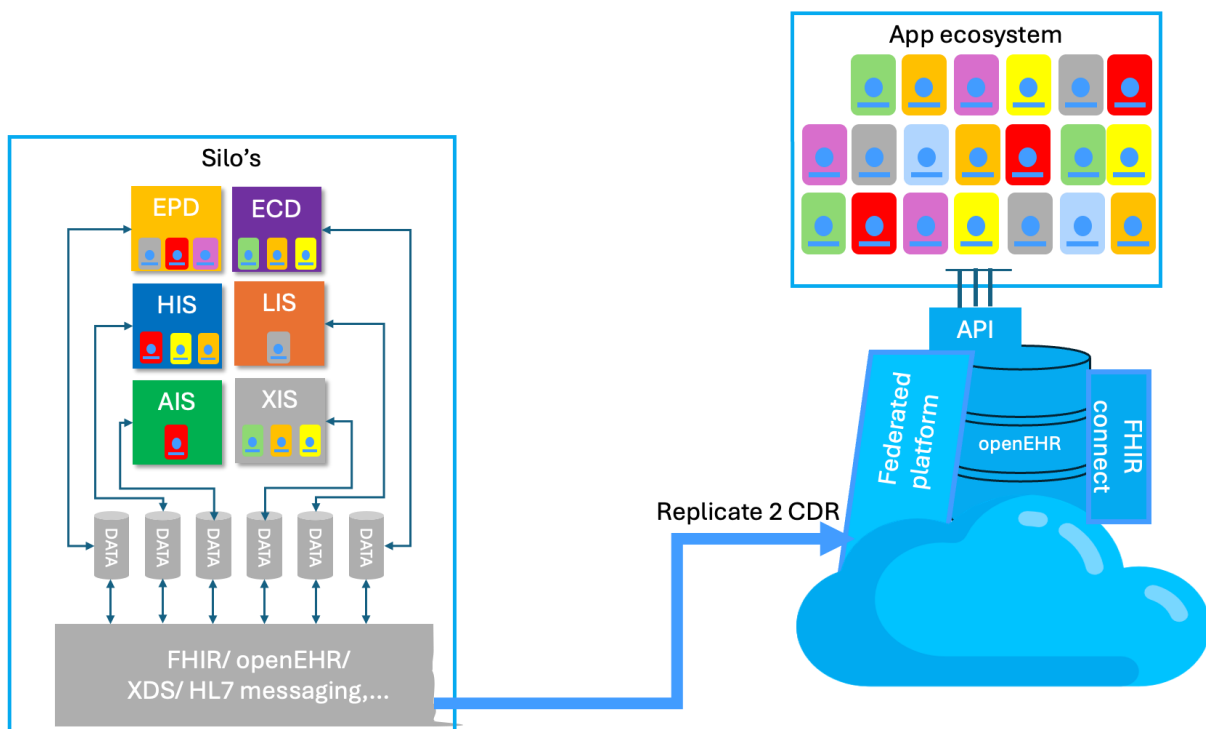
- Most silos have very limited capabilities for adding or updating data in a safe way, guaranteeing the integrity and validity of the added or updated data.

Replicate 2 CDR

The replicate2CDR flow will be used to replicate selected data from a subset of patients or clients, from the silo system to the CDR. RSO and its participants must decide on whose data will be replicated (does it include dead patients? Does it include patients whose last contact with the healthcare provider was longer than 10 years ago?). What data will be replicated depends on use case requirements and capabilities of silo systems.

This flow will involve:

- Utilizing every technology available for (real-time) synchronization of data from silos to the associated CDR. Possible methods include FHIR updates, unsolicited FHIR or HL7v2 messages, database polling, CDA documents, silo-specific APIs and even RPA.
- Once data is available in the 'vendor neutral' CDR, it is available for reuse in (future and unforeseen) use cases



Replication comes in two flavors: near real-time replication and batch replication. Near real-time replication means that data changes are applied as soon as they occur in the silo-system. Near real-time-replication depends on the ability of a silo-system to signal the HIE platform that data is added or updated. So called 'trigger events' in the silo

system (like admitting a patient or creating a laboratory result) trigger a signal to the HIE platform or trigger pushing information-updates to the HIE platform. Traditional healthcare technology to implement this type of replication is (HL7) messaging and XDS document registration. Modern variants include (FHIR) APIs on the HIE platform that can be used by the silo to push data, and event streaming. Batch replication on the other hand is based on scheduled tasks. The HIE platform regularly executes a replication task to fetch new and updated data from the silo. Technologies include scheduled API calls and database queries.

Technologies used for (near real-time) replication include those depicted in Figure 3:

- (Existing) messaging feeds
These include older but commonly used technologies such as HL7v2 messages and HL7 CDA documents, but also newer technologies like HL7 FHIR messages. These kind of messaging feeds are often formally supported by vendors.
- Realtime API data-feeds
Some silo systems are capable of calling outside APIs to send events or data based on some internal trigger. These kinds of technologies include event-streaming (combined with API calls) and unsolicited API calls that essentially push data-updates to the HIE platform. All kind of data formats could be utilized, mostly JSON or XML based. These include FHIR resources and openEHR compositions (for systems that are partly openEHR based such as Nedap or Code24)
- API query/ response
Many silo systems expose APIs that can be used to query for data, in respond to an event trigger (see real-time API data-feeds) or based on some scheduling mechanism. Exposed APIs may or may not be supported by vendors.
- Db query/ response
This means executing a database query in respond to an event trigger or based on some scheduling mechanism. Some systems expose stored procedures or database views for this purpose that are actively supported by vendors. Using these types of views and stored procedures are preferred to developing custom queries, because they tend to be more stable than the underlying physical table structure and are guaranteed (by the vendor) to produce correct results.
- XDS register-document
Some silos offer XDS capabilities and are able to register (on-demand) documents to an outside XDS registry. An HIE platform could use these

registrations to publish a reference to documents in the CDR.

- Robotic Process Automation (RPA). If all else fails, RPA or even human typing can come to the rescue.

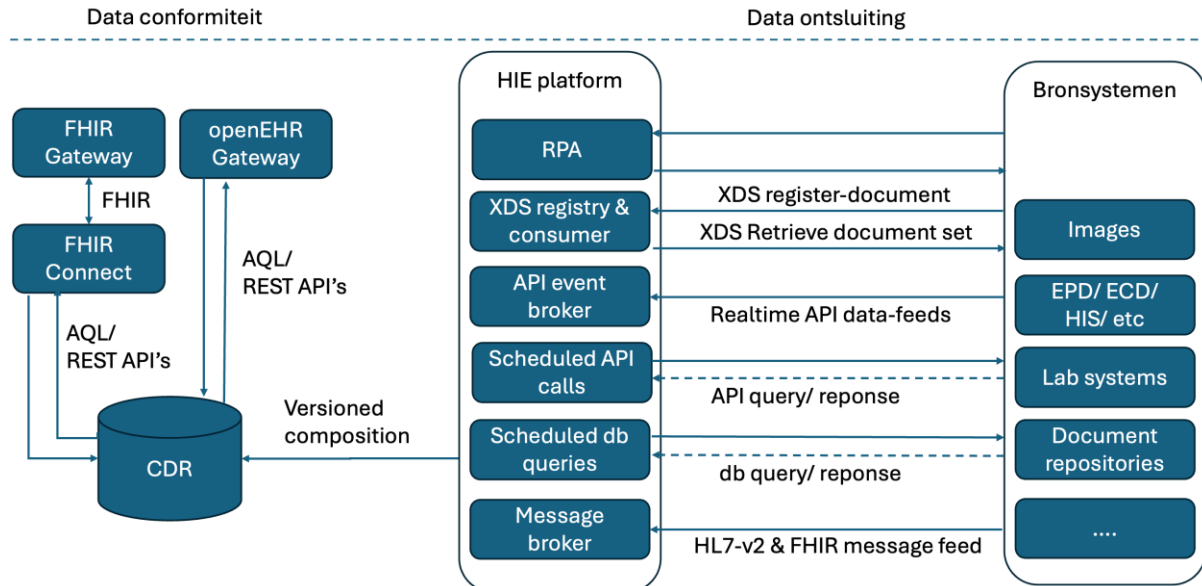


Figure 4: Technologies to replicate data in silos to a CDR

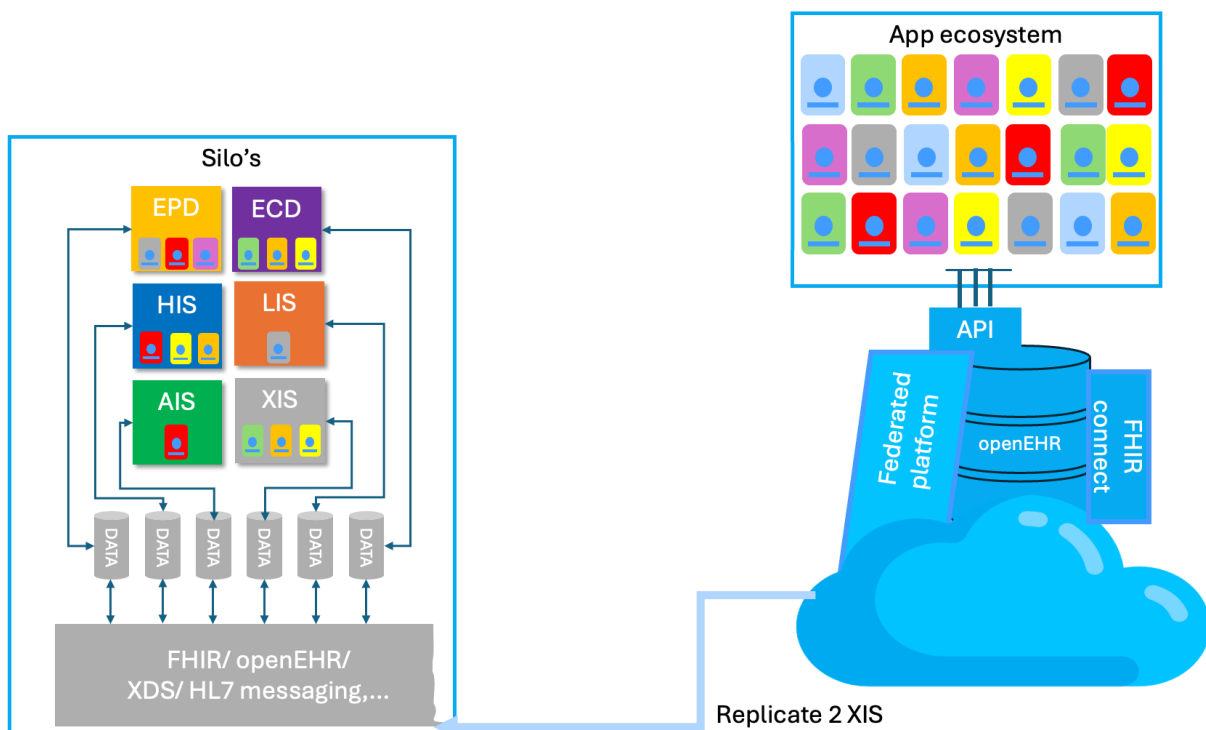
Challenges implementing this flow are:

- We are dealing with silos: options are very limited at times, and many options that are available are not (actively) supported by the silo vendors or only at high costs
- Software updates to silo-systems (some systems are updated very frequently) may break existing extractions and transformations, especially extractions that are not officially supported by the silo vendor
- Different silos require completely different technologies for extracting and transforming data. Even different versions of the same silo-system support- and/or require different extractions and transformations. There is no 'economy of scale' here, unless healthcare providers harmonize system-versions, share knowledge and align

Replicate from CDR

This flow will involve:

- Utilizing every technology available for (real-time) synchronization of data from the CDR to the associated silo(s). Possible methods include FHIR updates, unsolicited FHIR or HL7v2 messages, HL7v3 messages, CDA documents, database updates, silo-specific APIs or RPA.
- This flow will only be implemented when systems or processes break without it. In many cases viewing data using a context launch mechanism is sufficient, without requiring data to be added to or updated in the silo.



Challenges implementing this flow are:

- We are dealing with silos: options are very limited at times, and many options that are available are not (actively) supported by the silo vendors, or only at high costs
- Most vendors rule out non-supported methods for updating or adding data (contracts prohibit the use of such methods), because they might break the integrity of the silo.

- Software updates to silo-systems (some systems are updated very frequently) may break update mechanisms, especially when these mechanisms are not officially supported by the silo vendor
- Different silos require completely different technologies for extracting and transforming data. Even different versions of the same silo-system support- and/or require different technologies and data-formats. There is no 'economy of scale' here, unless healthcare providers harmonize system-versions

App Integration

The Zuid-Limburg data-ecosystem will eventually include a range of apps. Apps are integrated using one or more of the following options:

- A. Directly through the openEHR API layer of the federated platform. These apps can be used to read and write all data available in the platform.

This strategy provides the most complete and flexible method for integrating apps with the data-ecosystem. Many of these apps can be developed by citizen developers using the platform no-code/low-code tools. They can also be shared between different (inter)national programs and institutions, using the platforms marketplace.

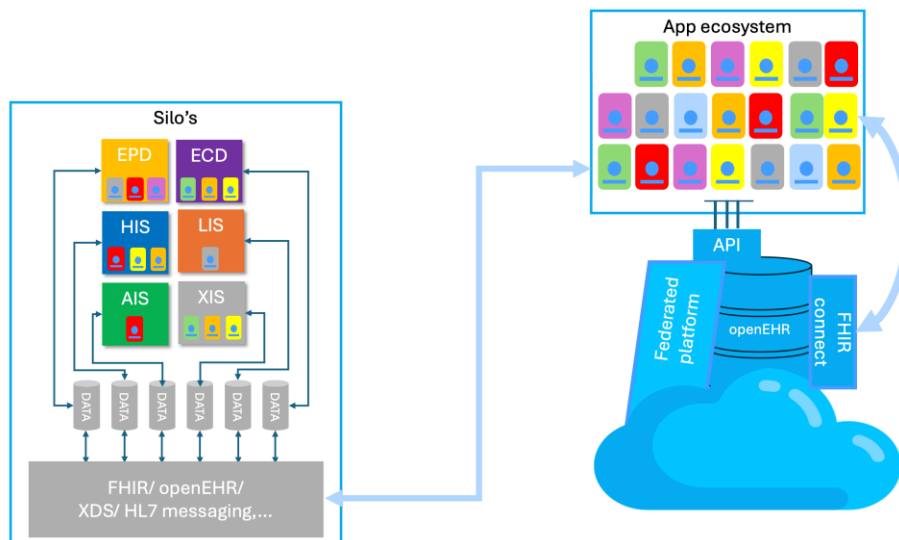
- B. Through the FHIR Façade of the federated platform, which is created using the openFHIR component and the FHIR Connect standard. These Apps can be used to read and write a smaller subset of the data available in the platform.

Although there are limitations on the (granularity of the) data that these apps can read and write, this strategy is still a flexible method for integrating apps with the data-ecosystem. The strategy opens the data-ecosystem for a wide range of existing apps that are based on the popular FHIR standard.

- C. Through the FHIR API's provided by silo systems. These apps can be used to read and write an even smaller subset directly from silo-systems.

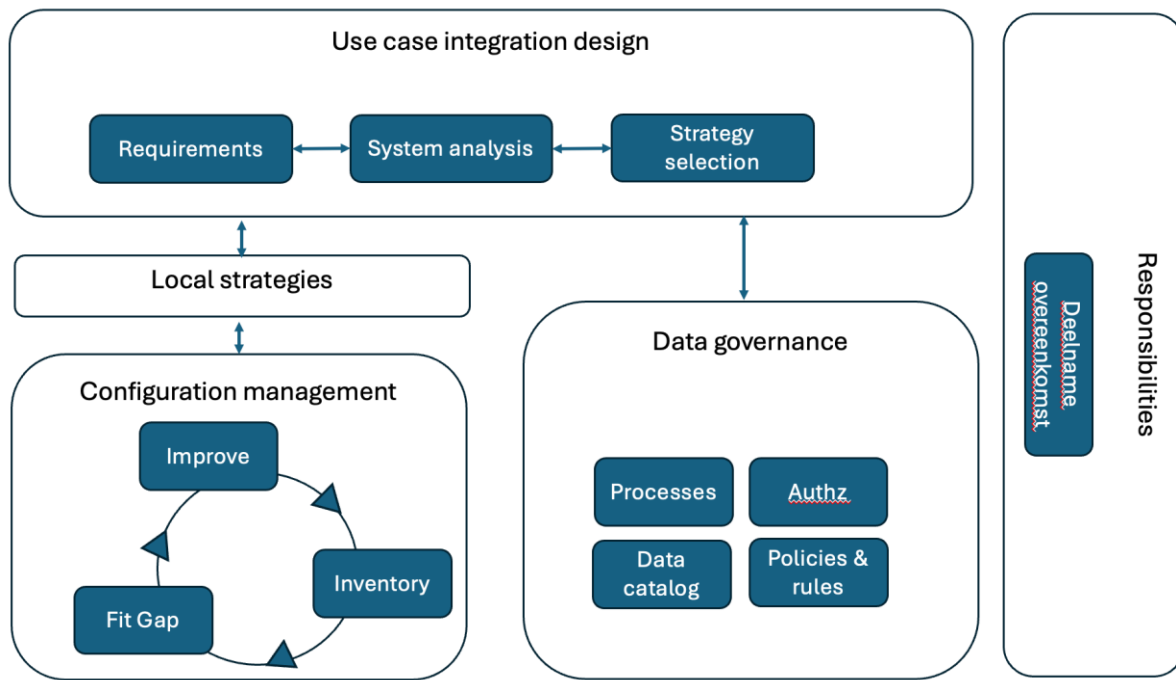
Although this is the least preferred option because it will require a (much) larger and repeating integration effort for each app that must be connected to the data-ecosystem, it can be a valid option in order to guarantee data integrity and a 'single source of truth'.

- D. Through a combination of the above



Local integration strategies

RSO participants are responsible for implementing the flows mentioned in 'Data flows' in and out of silo systems that are under their control and influence, such as EHR-systems and laboratory systems. Implementing these flows will require a mixture of integration strategies mentioned with each flow. Flows are implemented in a step-by-step fashion, and only as far as required for the use cases the participant participates in. RSO Zuid-Limburg will help participants choosing and implementing the most fitting integration strategies for each use case.



Responsibilities

The responsibilities that come with connecting a participant to the data-ecosystem will be laid down in an agreement between each participant and RSO Zuid-Limburg. Agreements are harmonized with the latest version of the TWIIN afsprakenstelsel, and involve topics like enforcing consent policies, implementing incident management procedures, implementing data governance and supporting agreed upon standards.

Regional configuration management

RSO Zuid-Limburg will setup regional configuration management for its participants, in order to have a clear and shared understanding of the integration capabilities of each participant and of the silo systems they control. This will include creating an inventory of all systems of record and the data they store (starting with BGZ elements), API's and extraction tools these systems provide and of the HIE platforms the participant uses to share data with other healthcare providers and patients.

The inventory will be used to:

- Discover gaps in the integration capabilities of each participant and suggest improvements
- Choose the best fitting combination of integration strategies to create a use case specific integration design for each use case the participant participates in

At the inventory (starting) phase of each transformation project, a dedicated team will

analyze the use case(s) and their data-/ integration needs. The configuration data will be extended with use case specific inventory for each participant.

Regional Data Governance

Healthcare professionals must be able to trust the data provided to them through the data-ecosystem. The ecosystems success, therefore, is highly dependent on the quality of that data. RSO Zuid-Limburg recognizes the following dimensions of data quality:

- **Correctness.**
Is the data correct? Does it not include flaws or errors?
- **Completeness**
Is the data complete (for example in relation to legislation and professional guidelines)
- **Consistency**
Is data consistent over different sources (like different health providers)
- **Validity**
Does the (stored or exchanged) data conform to relevant information standards?
- **Unicity**
Is the data unique over different sources (such as different systems and/ or different healthcare providers) or can it be deduplicated?
- **Timeliness**
Is data topical and not outdated at the time of use?

RSO Zuid-Limburg will setup regional data governance for its participants in order to raise data quality throughout the data-ecosystem. At the inventory (starting) phase of each transformation project, a dedicated team will analyze the use case(s) and their data-/ integration needs. The regional data governance board will select quality criteria, standards and other policies to apply to new data elements.

Use case integration design

At the inventory (starting) phase of each transformation project, a dedicated team will analyze the use case(s) and their data-/ integration needs. Data needs are analyzed and expressed using openEHR archetypes from the international CKM as much as possible.

The team will identify data sources and the differences between source data and the selected openEHR archetypes.

The team will identify the different integration strategies that can be used to fulfil data- and integration needs, based on the agreed upon responsibilities and the regional configuration management inventory. Together with each participant involved in the transformation project, the team will create a use case integration design that matches the capabilities of that participant and the requirements of the use case.

Vendor management

Unlike in the US and some other countries, no explicit policies for preventing vendors from information blocking strategies are in place in the Netherlands. Although Wegiz legislation forces healthcare providers to implement specific HIE for specific use cases, this is not preventing vendors from information blocking strategies based on pricing or imposing artificial barriers such as hiding or obfuscating necessary information and specifications. Also, modern solutions for sharing health- and healthcare data (such as vendor neutral storage) will lead to investments that some (or many) vendors are not willing or able to make. There are no stimulation policies in place for these kinds of major innovations in the Netherlands.

Overcoming information blocking will prove crucial for every major health data program, including the Zuid-Limburg data-ecosystem. RSO Zuid-Limburg will implement strategies for vendor management that include:

- Understanding and using the EHDS regulation and its potential impact on information blocking strategies.
- Influencing national policy makers and push for explicit policies that prevent information blocking by IT vendors. This will include pushing for modern technologies such as vendor neutral storage of health data, but also pushing for registration of available APIs, preferably in the Nictiz API library, and forcing vendors to implement generic APIs, such as the FHIR Questionnaire API, instead of use case specific APIs.
- Requiring (and helping) technology partners to publish API information in the Nictiz API library
- Working together with other regions, national programs and interest groups to harmonize requirements and customer demand, and share knowledge and solutions that help overcoming information blocking
- Help participants with the implementation of the 'Data flows'.