

Technical Support Instrument

Supporting reforms in 27 Member States

Joint Management of Data and Analysis for Swedish Healthcare Regions

*Sub project 4: Established conditions for a new national health
data platform*

2025-05-22



**Funded by
the European Union**



**Swedish Association
of Local Authorities
and Regions**

Sub-project 4

Established conditions for the development of regional health data platforms with a solution for transfer to a regionally shared platform

Tasks:

- Provide specifications for a visualisation area with user-friendly interfaces.
- Provide specifications for an analytical environment that meets different user needs.
- Provide specifications for functionality for the transmission of data from the platform to different stakeholders.
- Propose a data access solution that meets the requirements of quality assurance, privacy protection and efficiency.

Reference:

See Chapter 4, design of reference architecture



**Funded by
the European Union**



Strategi för regiongemensam vårddatahantering

Delprojekt 4

Joint Management of Data and Analysis for Swedish
Healthcare Regions

European Commission Directorate-General for Structural Reform Support

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Sammanfattning

Syfte med projektet

Vårddatahubben, baserad på en framtidssäkrad referensarkitektur, ska göra vårddata mer tillgänglig för sekundäranvändning, höja datakvalitet och nyttjandegrad, stödja uppföljning, analys, forskning, innovation och AI-tillämpningar samt katalysera regionernas arbete mot en mer datadriven hälso- och sjukvård generellt och i linje med EHDS-kraven.

Struktur & tillvägagångssätt

Arbetet i delprojektet har följt en strukturerad arbetsmetodik och en process med kontinuerlig förankring och diskussion med referensgrupper, experter inom relevanta områden, regioner och andra aktörer (tex. myndigheter).

Utvärdering av nuläge

Utvärderingen av nuläge har utgått ifrån resultatet från delprojekt 2 vad gäller regionernas dataförmågor, den tidigare framtagna datastrategin, en juridisk översikt samt benchlearning och diskussion med liknande initiativ. Exempel på dessa är HealthRI i Nederländerna, Genomic Medicine Sweden och SUSSA datalager.

Designkriterier

Designkriterierna är en första detaljering av de behov som den framtida lösningen ska tillgodose. De delas in i centrala områden som autentisering och auktorisering, datahantering, förvaltning och drift, juridik och organisation, rapportering och analys, säkerhet och behörighet, användning, integration, systemlogik och övrigt.

Arkitekturdesign



Arkitekturdesignen innehåller följande 5 delar. Introduktion och process som beskriver hur arbetet gått till, en referensarkitektur som är huvudleveransen, ett realiseringsexempel som konkretiserar referensarkitekturen, fördjupningsområden och summering av slutsatser från fördjupningsområden.

Fördjupningsområdena innehåller ett antal identifierande nyckelfrågor, tekniker och begrepp. Dessa fördjupningsområden innehåller en beskrivning om vad området är, varför det är relevant för datahubben och där det är relevant en slutsats.

Referensarkitektur

Referensarkitekturen består av fyra delar: Regional infrastruktur, Regiongemensam datahubb, Användning och Övergripande tjänster.

Inom regional infrastruktur utgår designen från befintlig infrastruktur och utökade förmågor. Referensarkitekturen nyttjar befintlig infrastruktur i den mån det går och ge förslag på behov av utökade förmågor (standardisering, mappning och lagring) för att säkerställa att regionerna behåller kontroll över sin data.

I den regiongemensamma datahubben ska data vara i ett standardiserat format, mappat till en gemensam datamodell och logiskt separerad. Data kan sen behandlas och mappas till nya ändamålsspecifika datamodeller (ex. för att skapa möjligheter för helt nya typer av analyser). I det sista skedet exponeras data via API eller delas via SFTP.

Användning tittar på hur ska data från datahubben användas.

Övergripande tjänster kallar vi den supportfunktionalitet som behövs för att hela referensarkitekturen ska fungera på ett effektivt och säkert sätt. Exempel på tjänster är pseudonymisering, datasäkerhet, metadatahantering och datakatalog.



Aktuella fördjupningsområden

Användarflöden
Dataintegration
Federerad eller central lagring
Datamodellering
Säkerhet
Pseudonymisering
Behörighetslösning
Metadatahantering och datakatalog
Existerande datainsamlingar
Cloud, hybrid eller on-prem
Säkra behandlingsmiljöer
Containrar

Analys av användningsfall

Analysen av användningsfall fokuserar på att beskriva olika scenarier där den regiongemensamma datahubben kan användas. Genom att analysera användningsfall kan projektet kunnat säkerställa att den framtida lösningen är anpassad till verkliga situationer.

Rekommendationer och nästa steg

Delprojektets rekommendationer har delats in i två arbetspaket:

- 1. Teknisk demonstration och kravställning** – fortsatt utreda, detaljera samt demonstrera delar av den föreslagna referensarkitekturen samt ingående komponenter
- 2. Teknisk realisering** – utreda och besluta om hur tekniska komponenter ska etableras och hur det påverkar regionernas organisation och arbetssätt, samt behov av investeringar.

**Nationellt system
för kunskapsstyrning
Hälso- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Innehåll

- 1 Tillvägagångssätt
- 2 Utvärdering av nuläge
- 3 Designkriterier
- 4 Design av arkitektur
- 5 Analys av användningsfall
- 6 Design av drift- och förvaltningsmodell
- 7 Rekommendationer och nästa steg

1

Tillvägagångssätt

Delprojekt 4 – Regiongemensam vårddatahubb

Delprojekt 4 utreder hur den regiongemensamma hubben för vårddata ska utformas. Detta innefattar såväl frågor kring arkitektur och funktionella och praktiska aspekter.

Syftet med delprojekt 4:

- Utforma en referensarkitektur som gör vårddata mer lättillgänglig för sekundäranvändning – mer effektivt och i större omfattning än idag.
- Vårddatahubben ska höja både datakvalitet och nyttjandegrad och stödja allt från uppföljning och analys till forskning, innovation och AI-tillämpningar.
- Med hubben som katalysator kan regionerna effektivisera sina resurser, höja ambitionen och sträva mot en mer datadriven hälso- och sjukvård.
- Referensarkitekturen ska dessutom vara framtidssäkrad och ge förutsättningar för fortsatt utveckling i linje med de kommande kraven inom EHDS.



Struktur för arbetet i delprojektet



**Nationellt system
för kunskapsstyrning
Hälso- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Deltagare i processen (1/2)

Delprojektmedlemmar

- Elisabeth Berglön, Sjukvårdsregion Mellansverige
- Catarina Karlberg, Västra sjukvårdsregionen
- Johan Tollebrant, Sjukvårdsregion Stockholm-Gotland
- Rolf Rönnback, Chefsarkitekt på Inera
- Peder Hofman-Bang, Projektledare SKR
- Alexander Rosén, Expert inom hälsosektorn EY
- Madelene Rundin Geuken, Projektledare EY
- Jacob Sandefeldt, Lösningssarkitekt EY
- Alice Paulsen, Projektledarstöd EY
- Erik Vermeulen, International expert inom hälsodata EY
- Hanna Pohjonen, Internationell expert inom hälsodata

Nationell arbetsgrupp, NAG

Sjukvårdsregionernas representanter:

- Arvid Widenlou Nordmark, Norra sjukvårdsregionen
- Elisabeth Berglön, Sjukvårdsregion Mellansverige
- Johan Tollebrant, Sjukvårdsregion Stockholm-Gotland
- Åsa Berling, Södra sjukvårdsregionen
- Reidar Källström, Sydöstra sjukvårdsregionen
- Catarina Karlberg, Västra sjukvårdsregionen

Medlemmar arbetsgrupp arkitektur

- Mats Jogbäck, Arkitekt SUSSA Data och Analys (SDA), Region Västerbotten
- Jonas Osterman, Arkitekt SUSSA Data och Analys (SDA), Region Västerbotten
- Maria Holmberg, Arkitekt SUSSA Data och Analys (SDA), Region Sörmland
- Erik Frumerie, Arkitekt, Västra Götalandsregionen
- Hayder Al-Zubaidi, Arkitekt, Västra Götalandsregionen

Nationell samverkansgrupp data och analys, NSG HD (Styrgrupp)

Ordförande

- Åsa Dederling, hälso- och sjukvårdsdirektör, Region Dalarna, Sjukvårdsregion Mellansverige

Ledamöter

- Anna Granevärn, tillförordnad hälso- och sjukvårdsdirektör, Region Härjedalen, Norra sjukvårdsregionen
- Åsa Dederling, hälso- och sjukvårdsdirektör, Region Dalarna, Sjukvårdsregion Mellansverige
- Anders Ahlsson, hälso- och sjukvårdsdirektör, Region Stockholm, Sjukvårdsregion Stockholm-Gotland
- Ann-Marie Schaffrath, digitaliseringsdirektör, Västra Götalandsregionen, Västra sjukvårdsregionen
- Annica Öhrn, hälso- och sjukvårdsdirektör, Region Östergötland, Sydöstra sjukvårdsregionen
- Jörgen Wenner, biträdande hälso- och sjukvårdsdirektör, Region Skåne, Södra sjukvårdsregionen
- Tobias Dahlström, representant för ekonomidirektörsnätverket, ekonomidirektör, Region Dalarna
- Helén Lundkvist Nymansson, chef för sektionen för data och analys, SKR
- Mari Forslund, SKR (tillfälligt i väntan på kommunrepresentant)

Adjungerad

- Johannes Hörnberg, adjungerad styrgruppsmedlem från Digitaliseringsnätverket, Region Jämtland Härjedalen

**Nationellt system
för kunskapsstyrning
Hälso- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Deltagare i processen (2/2)

Utöver projektets medlemmar och arbetsgrupp har ett flertal personer varit engagerade i delprojektet i egenskap av referensgrupp, expertgrupp samt genom benchlearning från andra initiativ.

Namn	Organisation
Anna Dovärn	Socialstyrelsen
Ann Gustafsson	E-hälsomyndigheten
Carla Alvarez Costa	Katalonien
Carlo Berg	GENSAM
Christian Högberg	Edence health (OMOP)
Christian Johansen	Biobanksregistret (Biobank Sverige)
Fredrik Quistgaard	Vetenskapsrådet, RUT, Senash
Hanna Douglas Thomee	Region Halland, Arkitekturrådet
Helene Spjuth	Region Stockholm, Centrum för hälsodata
Henrik Vennersten	IT-arkitekt, Region Halland
Joel Hedlund	Linköpings Universitet
Jonas Nilsson	Datasäkerhet, SKR
Leone Flikweert	CEO, Health-RI
Lovisa Sydén	Folkhälsomyndigheten
Manne Andersson	E-hälsomyndigheten, Samverkansarkitektur
Manólis Nymark	Jurist, SKR

Namn	Organisation
Maria Nilsson	Vetenskapsrådet
Markus Lingman	Region Halland
Michael Peolsson	E-hälsomyndigheten
Mårten Mannström	Region Stockholm, Arkitekturrådet
Niklas Eklöf	Socialstyrelsen (dubbelkolla om han varit med)
Niels Bolding	Health-RI
Olof Elfwerig	SCB, MONA
Per Sikora	Genomic Medicine Sweden
Per Wiklander	Vetenskapsrådet
Peter Hedman	Biobanksregistret (Biobank Sverige)
Saif Al-Mobarek	SUSSA Data och Analys (SDA)
Susan Sverin	Socialstyrelsen
Susanne Dahllöf	SCB, MONA
Tomas Snäckerström	Biobanksregistret (Biobank Sverige)
Torbjörn Eles	IT-ansvarig på Inca (nationell IT-plattform Cancer)

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Personerna har verkat rådgivande under projektets gång men kan ej ses som ansvariga för dess resultat

2

Kartläggning av nuläge

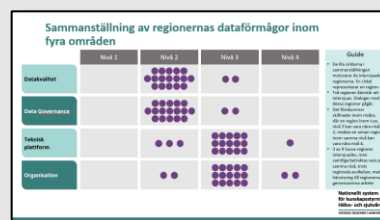
Kartläggning av nuläge lade grund för designkriterier

Projektets kartläggning av nuläge innefattar mognadsanalys av regionerna, kartläggning av juridiska förutsättningar samt best practices från liknande initiativ*

Regionernas dataförmågor (Delprojekt 2)

Som en del i analysfasen för delprojekt 2 genomfördes en sammanställning över dataförmågor inom fyra områden.

Resultat från de intervjuade regionerna sammanställdes och analyserades för att skatta mognadsnivån och utforma rekommendationer kring hur regionerna kan förbättra sina dataförmågor.



Strategi för gemensam vårddatahantering

2022 genomfördes ett arbete med att ta fram en gemensam datastrategi för regionerna. Delar i strategin har använts som input till hubbens designkriterier.



Juridiska förutsättningar



Juridiska förutsättningar kartlades, och de lagar med störst bedömd inverkan på projektet identifierades.



Ett appendix** kring regulatorisk översikt utvecklades, med utgångspunkt i material från arbetet med strategi för gemensam vårddatahantering.



Nästa steg kopplat till juridik identifierades. Designkriterier med juridiska överväganden ingår i projektet medan rättslig analys och dataskyddskonsekvensbedömning är utanför scope.

Liknande initiativ

Projektet har identifierat och analyserat best practices från fem initiativ som har likheter med detta projekt. Initiativen gav även indikationer kring finansiering samt drift- och förvaltningsmodell.



Genomic Medicine Sweden



Sussa datalager



RUB



Health-RI



Katalonien

Dessa delar utgjorde grund för framtagande av **designkriterier** för vad den framtida vårddatahubben ska åstadkomma

Nationellt system för kunskapsstyrning Hälso- och sjukvård

SVERIGES REGIONER I SAMVERKAN

* Mer detaljer finns i [Appendix 1](#) ** Se [Appendix 3](#)

3

Designkriterier

Designkriterier är en första detaljering av de behov som den framtida lösningen ska tillgodose

Designkriterierna besvarar frågeställningen: *Vad ska uppnås med den framtida vårddatahubben?*



Typer av designkriterier

Designkriterier delas vanligtvis in i ett antal centrala områden som underlättar utformning av den framtida referensarkitekturen.

Designkriterier kan antingen vara "funktionella designkriterier" samt "icke-funktionella designkriterier"

Funktionella designkriterier

definierar **vad** den framtida lösningen ska åstadkomma och möjliggöra – exempel: "Den framtida vårddatahubben ska möjliggöra regionalt tillgängliggörande av hälsodata"

Icke-funktionella designkriterier

definierar **hur och på vilket sätt** den framtida vårddatahubben ska utföra efterfrågad funktionalitet

Delprojektet har:

- Prioriterat de funktionella designkriterierna
- Benämnt samtliga övergripande designkriterier som "ska-kriterier" (dvs exkluderat bör-kriterier)
- Arbetat fram 56 övergripande designkriterier för den framtida hubben*

Designkriterier för ny vårddatahubb

Designkriterierna har delats in i 10 designområden



Beskrivning av tre områden*

1. Användning	7. Rapportering och analys
<p>I området ingår designkriterier som handlar om att stödja både nuvarande och framtida användarfall, främja datadrivet beslutsfattande och resurseffektivisering.</p> <p>Kriterierna betonar vikten av skalbarhet, intuitivitet och samarbete. Dessutom fokuserar de på att underlätta delning av hälsodata och bidra till kunskapshöjning inom vården.</p>	<p>I området ingår designkriterier som handlar om att möjliggöra en säker och användarvänlig lösning för utdatahantering.</p> <p>Kriterierna fokuserar på att erbjuda ett modernt gränssnitt för analys av hälsodata oavsett regionalt ursprung och att stödja både deskriptiv och prediktiv analys. Dessutom ska infrastrukturen kunna hantera stora datamängder och komplexa analyser.</p>
3. Datahantering, datakvalitet och data management	
<p>I området ingår designkriterier som handlar om att använda internationella och leverantörsneutrala standarder för att säkerställa datakvalitet och interoperabilitet.</p> <p>Kriterierna fokuserar på att möjliggöra kontinuerlig datahantering och automatiserade valideringsmekanismer, samt att definiera tydliga roller och ansvar för dataäggande och delning. Spårbarhet av dataflöden, versionshantering av dataset och inbyggd kvalitetssäkring av inkommande data är också del i designkriterierna.</p>	

**Nationellt system
för kunskapsstyrning
Häls- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Beskrivning av samtliga områden finns i [Appendix 1](#)

4

Design av arkitektur

Översikt över innehåll i avsnitt 4: Design av arkitektur

Följande avsnitt är utformat enligt nedan struktur.		
	Introduktion och process	Innefattar introduktion till avsnittet samt processen som lagt grund för referensarkitekturen.
	Referensarkitektur	Beskriver referensarkitekturen utifrån regional infrastruktur, region-gemensam datahubb, övergripande tjänster och användning.
	Realiseringsexempel	Beskriver hur referensarkitekturen kan realiseras.
	Fördjupningsområden	Beskriver referensarkitekturens komponenter mer detaljerat.
	Summering slutsatser	Summerar slutsatserna från fördjupningsområdena.

Aktuella fördjupningsområden	Sidor
Användarflöden	29-34
Dataintegrering	35-36
Federerad eller central lagring	37-46
Datamodellering	47-51
Säkerhet	52-53
Pseudonymisering	54-58
Behörighetslösning	59-61
Metadatahantering och datakatalog	62-67
Existerande datainsamlingar	68-69
Cloud, hybrid eller on-prem	70-72
Säkra behandlingsmiljöer	73-75
Containrar	76-78

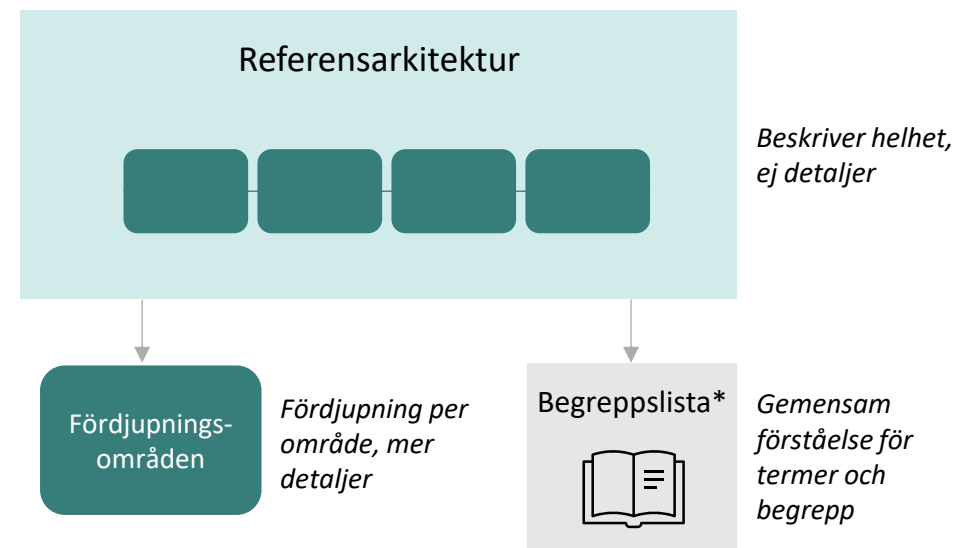
Den övergripande referensarkitekturen kan detaljeras i ett antal fördjupningsområden

Referensarkitektur

- En referensarkitektur kommunicerar en målbild av ett fysiskt system för en specifik lösning. Den beskriver komponenterna i lösningen samt hur de ska interagera.
- Syftet är att ge stöd och vägledning vid utveckling av lösningar, förbättra kommunikation mellan intressenter och säkerställa att alla parter har en gemensam förståelse för systemets struktur och funktionalitet.
- En referensarkitektur kan också fungera som en grund för att jämföra olika lösningar och teknologier, vilket underlättar beslut om vilka komponenter som ska användas.
- Sammanfattningsvis är en referensarkitektur ett viktigt verktyg för att strukturera och effektivisera utvecklingsprocessen.

Fördjupningsområden

- Referensarkitekturen kan utvecklas inom flera fördjupningsområden, där varje område analyseras djupare och beskrivs mer detaljerat.
- Fördjupningsområdena är i regel de delar av arkitekturen som är mest kritiska för funktionen, alternativt svårast att realisera, eller områden där man måste göra val utifrån lokala omständigheter.



* En lista över relevanta termer och begrepp samt en beskrivning av dessa finns i [Appendix 1](#)



Process vid framtagande av referensarkitektur

Den framtagna referensarkitekturen har baserats på kunskap från flertalet källor och erfarenheter från andra initiativ, samt förankrats och validerats med referensgrupper och utvalda experter. Processen är att betrakta som iterativ, där aktiviteter pågått parallellt och i cykler.

1 Etablera arbetsgrupp

- En arbetsgrupp bestående av arkitekter från Sussa, VGR, EY och Inera.
- Beslut kring arbetssätt baserat på förutsättningar kring tillgänglighet.



2 Definiera fördjupningsområden

- Skapa lista över områden/nyckelkoncept för referensarkitekturen.
- Prioritera områden.

3 Informationsinhämtning

- Arbeta med fördjupningsområden enligt prioritering.
- Benchlearning från liknande initiativ.
- Identifiera relevanta informationskällor.
- Detaljera förslag.



4 Diskussion och slutsats

- Diskussion med arbetsgrupp och andra för området relevanta parter.
- Slutsats kring området.



5 Iterera fördjupningsområden

- Validera och eventuellt uppdatera fördjupningsområdenas fortsatta relevans och prioritet baserat på resultat från diskussion och slutsats.

6 Skapa referensarkitektur

- Skapa referensarkitektur.
- Uppdateras efter resultat från arbetet med fördjupningsområden.

7 Användningsfall

- Applicera den framtagna referensarkitekturen på användningsfall från delprojekt 1 för att testa referensarkitekturen.

8 Validering och Förankring

- Förankring med referensgrupp.
- Förankring med myndigheter (Socialstyrelsen, E-hälsomyndigheten, Vetenskapsrådet, SCB med flera).



9 Dokumentation

- Dokumentera resultat och slutsats.
- Rekommendationer kring nästa steg.

Referensarkitektur



**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN



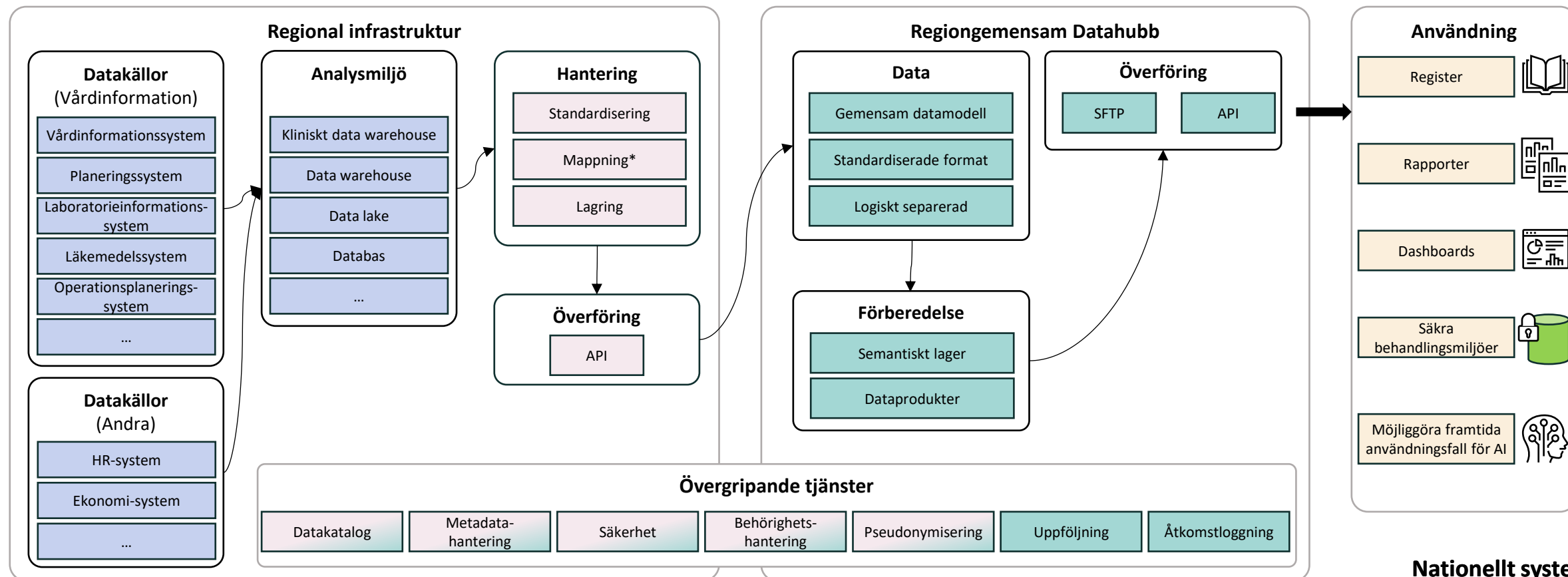
4.1 Referensarkitektur



Referensarkitektur

Referensarkitekturen består av 4 delar: Regional infrastruktur, Regiongemensam datahubb, Användning och Övergripande tjänster. Syftet och påverkan av de olika delarna beskrivs i följande bilder. Denna bild ger en överblick över hela flödet genom vårddatahubben.

- Befintlig regional funktionalitet
- Utökad regional funktionalitet
- Regiongemensam funktionalitet
- Användning



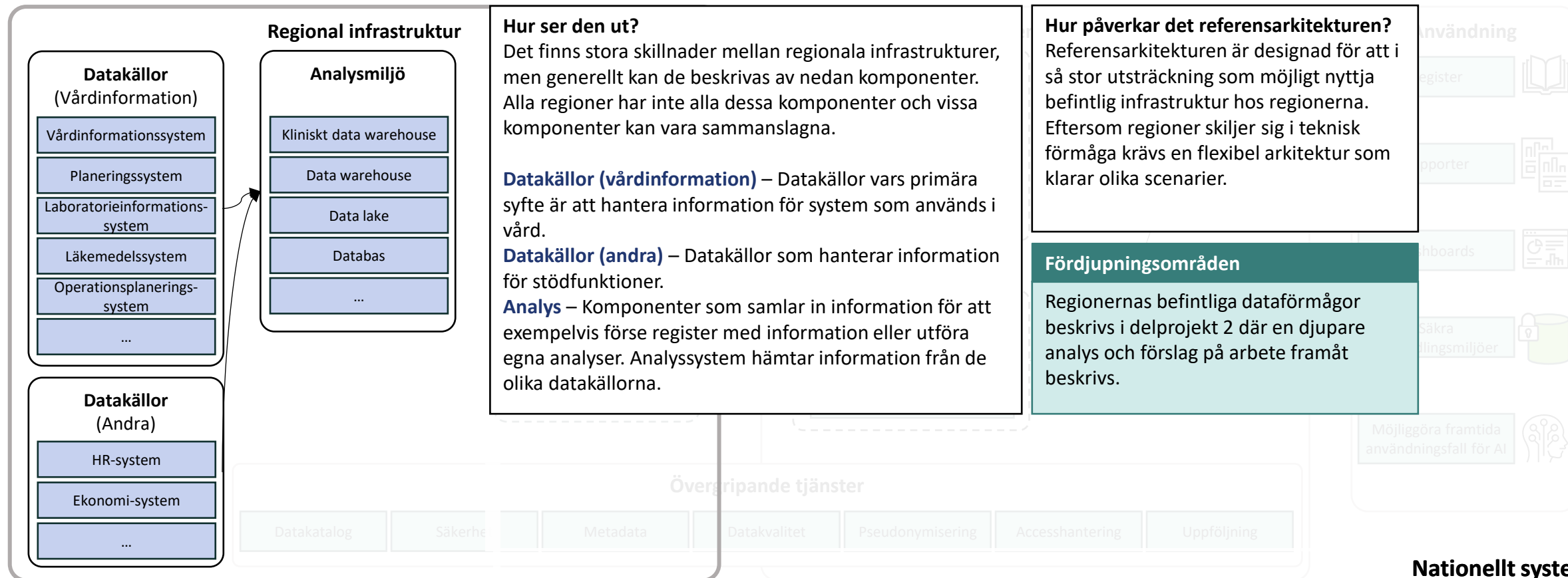
**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Mappning från regional modell/modeller till regiongemensam modell

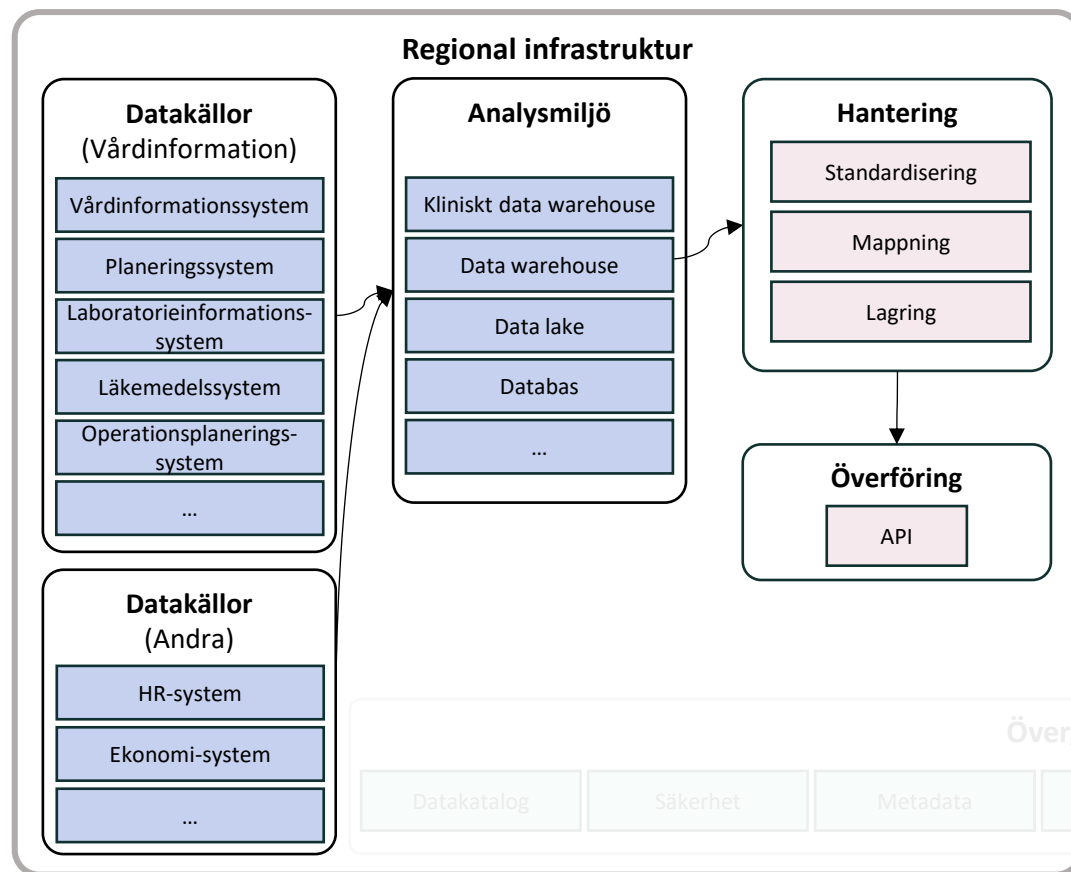


Regional infrastruktur – befintlig





Regional infrastruktur – utökning



Varför behövs utökad regional funktionalitet?

En av de viktigaste principerna för att realisera värdet av en regiongemensam datahubb är att data måste vara jämförbar mellan regioner.

Vilken funktionalitet kommer behövas?

För att säkerställa att data är jämförbar kommer en regiongemensam datamodell och ett enhetligt lagringsformat per informationsmängd behövas. För att hantering av mappning till den regiongemensamma datamodellen ska vara hållbar över tid, och för att kontroll och ägandeskap ska stanna inom den egna regionen, kommer regionerna behöva vara ansvariga för att mappa data från sina befintliga lagringsmodeller till den regiongemensamma. De behöver också standardisera format.

Data som regionerna modellerat kommer i vissa fall behöva pseudonymiseras* och lagras beroende på användningsfall och befintliga förmågor.

Det sista som sker är att data överförs från de regionala miljöerna till den regiongemensamma datahubben via API:er som kontrolleras av regionen.

Fördjupningsområden

Följande områden ingår i eller kompletterar referensarkitekturen och kommer utredas i separata sektioner:

- Datamodellering
- Pseudonymisering
- Dataintegrering**
- Containerar***

Möjliggöra framtida användningsfall för AI



**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Övergripande tjänst

** Förflyttning av data mellan regionala infrastrukturer och regiongemensam datahubb

*** Potentiell realisering av utökad regional funktionalitet



Regiongemensam datahubb

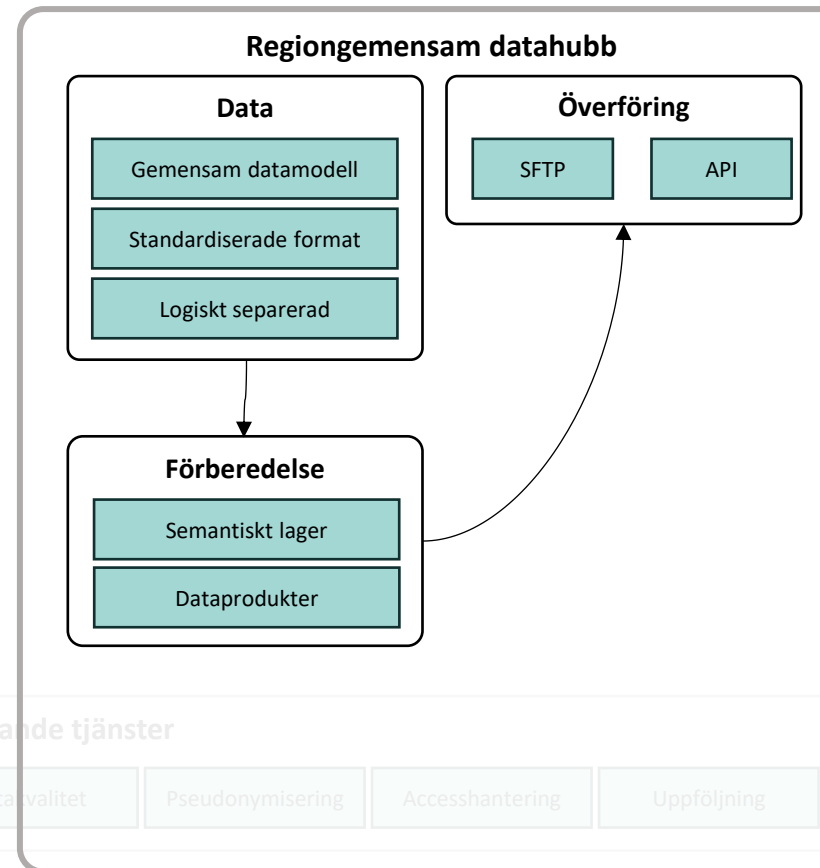
Om Regiongemensam datahubb

Idag skickas data från regionerna till massa olika aktörer i flera olika syften. Ofta är det samma eller likartade datamängder som skickas. I framtiden kan den regiongemensamma datahubben vara en väg in för regionerna att dela data och en väg ut för konsumenter att få ut data.

Hur ska en Regiongemensam datahubb möta syftet?

I den regiongemensamma datahubben samlas data från olika regioner i ett enhetligt standardiserat format i en gemensam datamodell. Data kommer fortfarande vara logiskt separerad för att data av juridiska skäl inte får blandas mellan regioner utan prövning.

I datahubben skapas ändamålsenliga dataprodukter där en delmängd av data modelleras mot ett specifikt syfte och sedan görs redo för överföring. Ett exempel är dataöverföring till ett specifikt register där enbart en delmängd av tillgänglig data behövs för att tillgodose registrets krav.



Fördjupningsområden

Följande områden ingår i eller kompletterar referensarkitekturen och kommer utredas i separata sektioner:

- Dataintegrering*
- Central lagring eller federering**
- Datamodellering

Möjliggöra framtida användningsfall för AI



**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Förflyttning av data mellan regionala infrastrukturer och regiongemensam datahubb

** Välgval gällande realisering av regiongemensam datahubb



Övergripande tjänster

Vad är Övergripande tjänster?

Övergripande tjänster är ett samlingsbegrepp som innefattar plattformsövergripande tjänster för hela vårddatahubben och tjänster som täcker både vårddatahubben och de regionala infrastrukturerna.

Fördjupningsområden

Följande områden ingår i eller kompletterar referensarkitekturen och kommer utredas i separata sektioner:

- Metadatahantering
- Datakatalog
- Säkerhet
- Behörighetshantering
- Pseudonymisering

Varför är dessa koncept Övergripande tjänster?

Datakataloger kommer behöva skapas på regional nivå och sedan antingen federeras eller synkroniseras regiongemensamt för att tillgänglig data i datahubben ska vara dokumenterad och sökbar.

Metadatahantering är arbetet som möjliggör datakataloger.

Säkerhet avser hur data hanteras i vila, transport och användning, både i den regionala infrastrukturerna, i vårddatahubben och däremellan.

Behörighetshantering styr tillgång till data. Exempelvis, eftersom regionerna har tillgång till sin egen data i datahubben kommer åtkomster behöva synkroniseras mellan datahubben och de regionala infrastrukturerna.

Pseudonymisering innebär att byta ut identifierbar information (namn, personnummer, etc.) mot artificiella identifierare eller pseudonymer. Syftet är att göra det svårare att koppla samman attribut till en individ utan ytterligare information. Eftersom pseudonymisering kan ske på både regional och regiongemensam nivå behöver tjänsten vara övergripande.

Uppföljning och **Loggning** syftar till att möjliggöra uppföljning för bland annat revision och compliance.

Övergripande tjänster

Datakatalog

Metadata-
hantering

Säkerhet

Behörighets-
hantering

Pseudonymisering

Uppföljning

Åtkomstloggning

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN



Användning

Om Användning

Referensarkitekturen har skapats för att kunna möta användningsfall som krävs av respektive användningsmönster (Register, Rapporter o.s.v.). Dessa användningsfall har sen grupperats in i 4 olika användarflöden. Dessa användarflöden är vad referensarkitekturen i första hand ska möjliggöra.

Fördjupningsområden

Följande områden ingår i eller kompletterar referensarkitekturen och kommer utredas i separata sektioner:

- Användarflöden
- Nationella register och kvalitetsregister
- Säkra behandlingsmiljöer

Användning

Register



Rapporter



Dashboards



Säkra
behandlingsmiljöer



Möjliggöra framtida
användningsfall för AI



**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

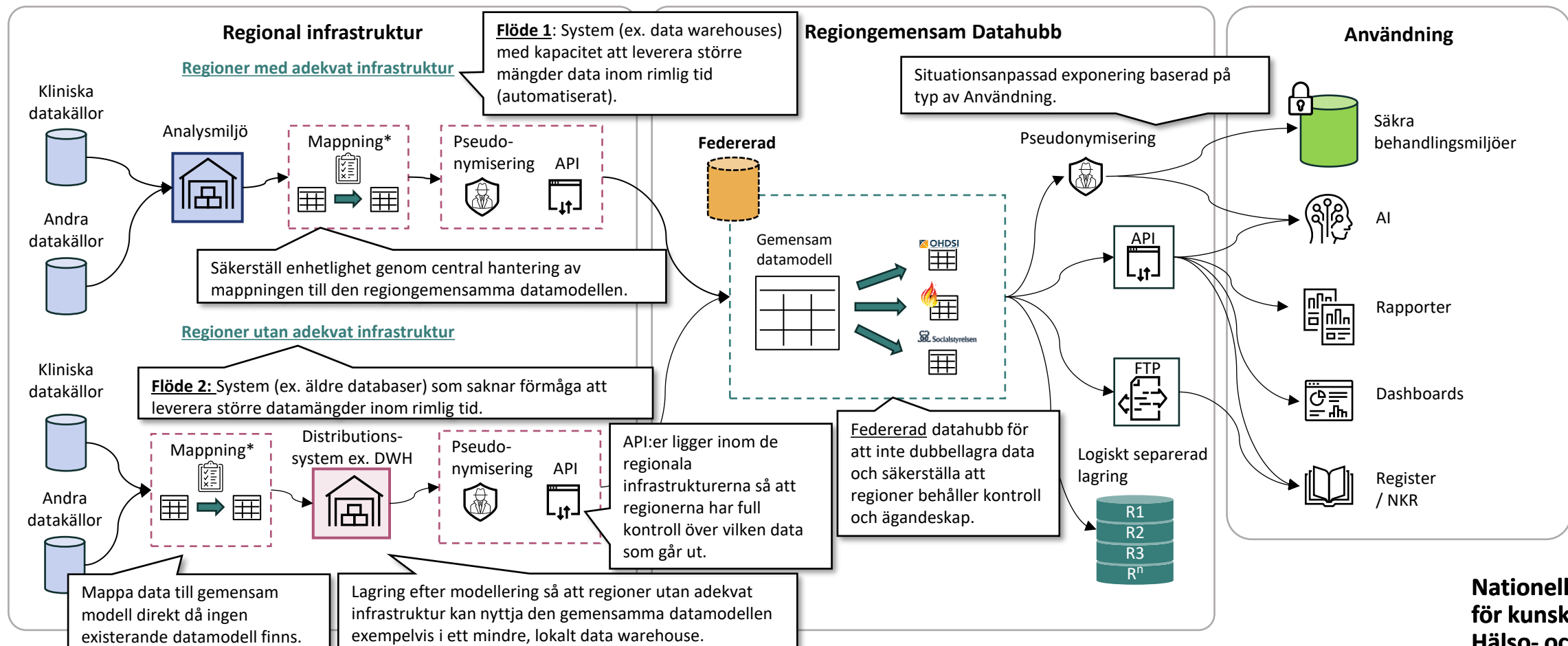


4.2 Realiseringsexempel

Realiseringsexempel

Eftersom regionerna har varierande mognadsgrad för sina regionala infrastrukturer kommer arkitekturen behöva vara flexibel och kunna anpassas efter olika regionala förutsättningar. Här exemplifieras två separata flöden för att ta hänsyn till olika regionala förutsättningar.

- Befintlig regional funktionalitet
- Utökad regional funktionalitet
- Regiongemensam funktionalitet
- Användning



* Mappning från regional modell/modeller till regiongemensam modell

4.3 Fördjupningsområden

Användarflöden

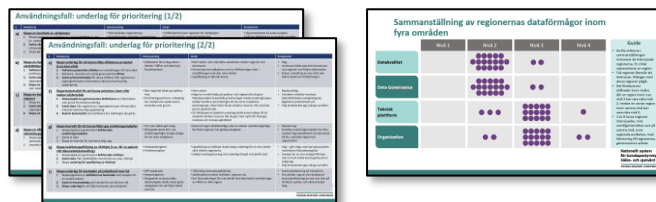
Användarflöden – ett verktyg för att designa en referensarkitektur

Om användarflöden

Ett användarflöde är en gruppering av användningsfall som har använts för att identifiera de primära användningsområdena för den regiongemensamma datahubben. Användarflödena är övergripande eftersom många olika scenarier och typer av användningsfall ska hanteras i datahubben.

Hur har de tagits fram?

Användarflödena har tagits fram genom omvärldsbevakning, intervjuer från regionerna i delprojekt 2 och de framtagna användningsfallen i delprojekt 1.



Hur har de använts?

Genom att analysera användarflöden har vi identifierat nödvändiga komponenter för att möta regionernas behov, samtidigt som vi säkerställer att framtida funktionalitet, inklusive AI-relaterade möjligheter, inte begränsas.

Identifierade användarflöden

Analys över en eller flera regioner där användaren inte är dataägare

En region vill göra analyser på sin egen data

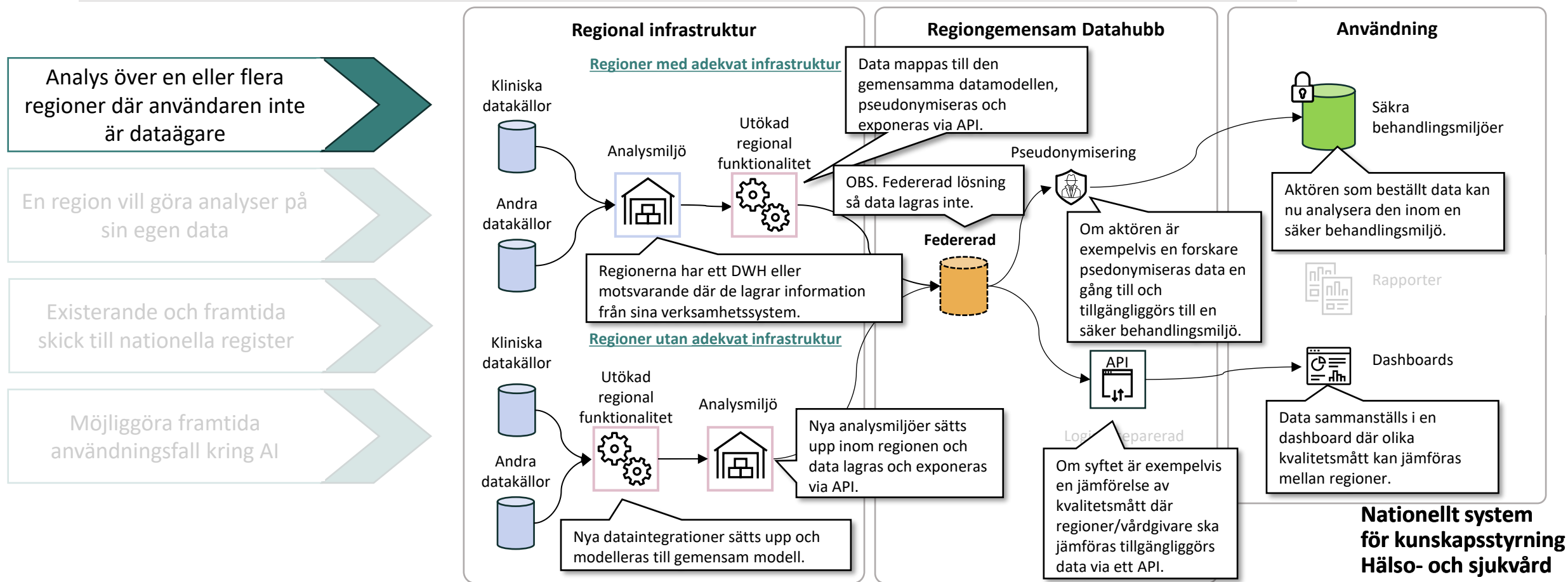
Existerande och framtida skick till nationella register

Möjliggöra framtida användningsfall kring AI

Användarflöde 1 - exempel

En förenklad bild av realiseringsexemplet används för att exemplifiera användarflöden. Exempel på hur de olika användarflödena skulle kunna realiseras kommer redogöras för i följande bilder. I användarflöde 1 beskrivs ett scenario där en förfrågan har **kommit in, behandlats och godkänts** för att lämna ut data.

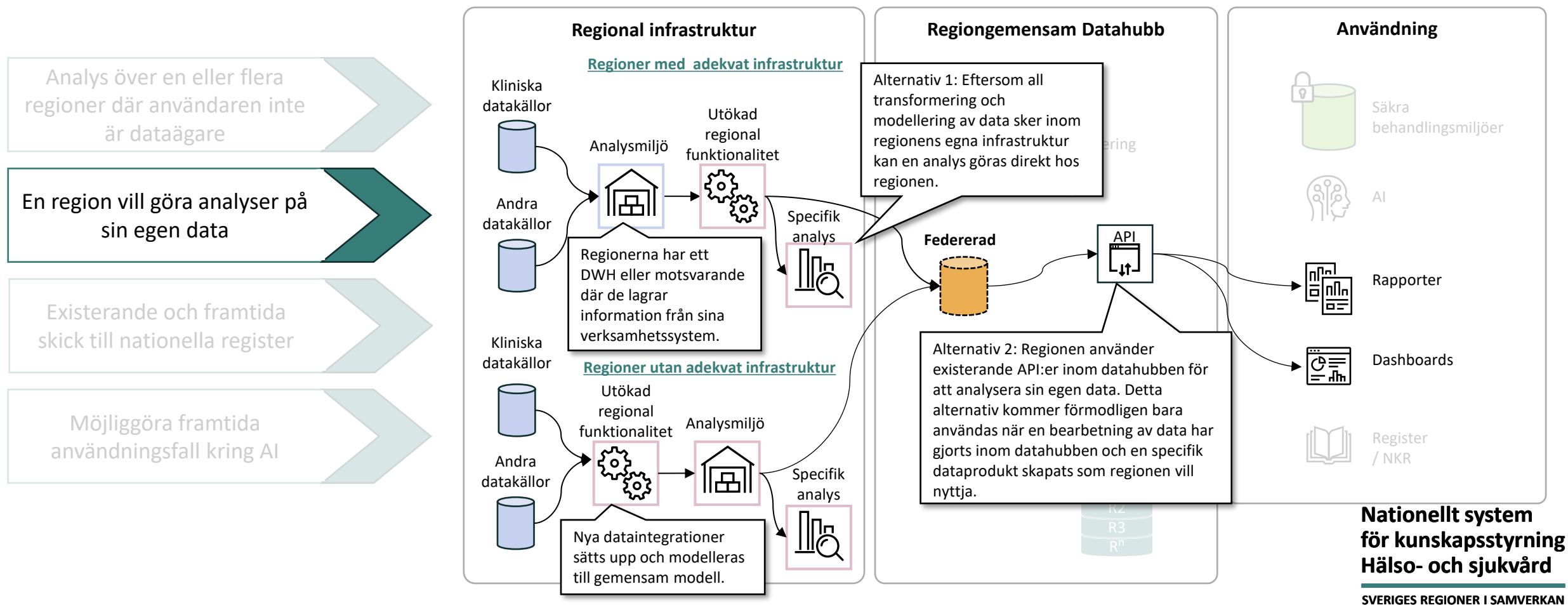
- Befintlig regional funktionalitet
- Utökad regional funktionalitet
- Regiongemensam funktionalitet
- Användning



Användarflöde 2 - exempel

I användarflöde 2 har **ingen förfrågan** inkommit utan regionen som dataägare vill utföra en analys på en datamängd som tillgängliggjorts i den regiongemensamma datahubben.

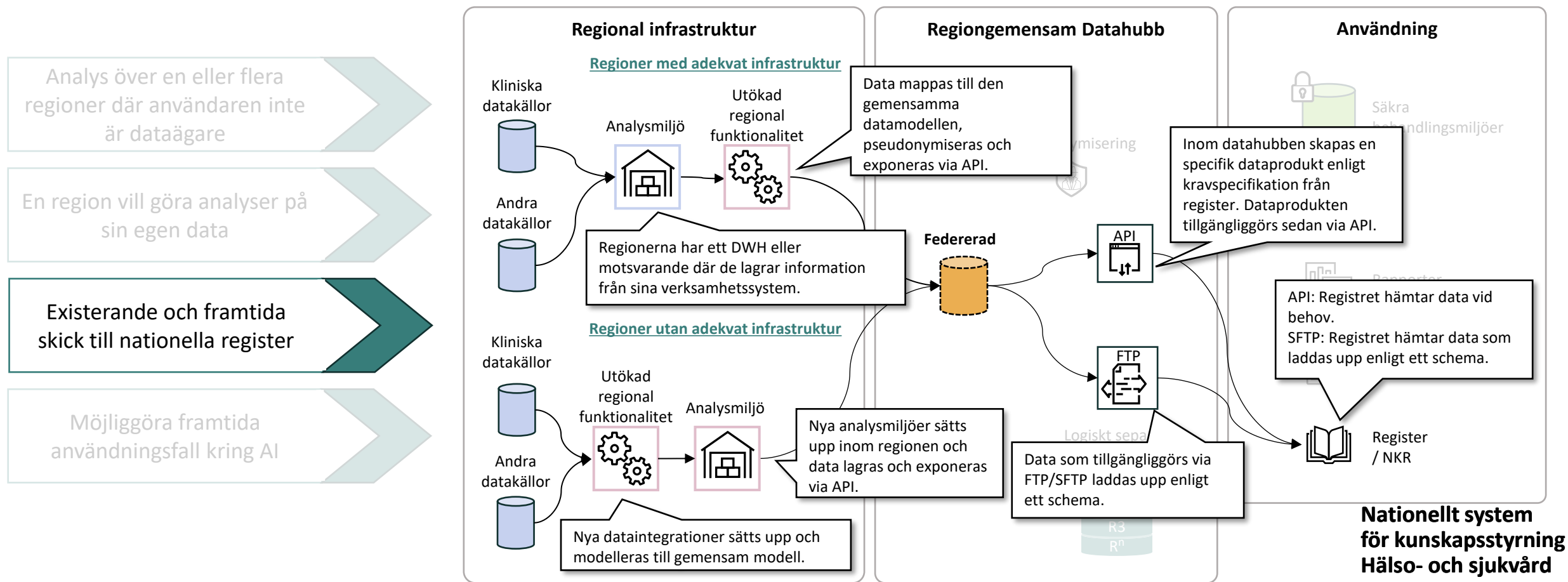
- Befintlig regional funktionalitet
- Utökad regional funktionalitet
- Regiongemensam funktionalitet
- Användning



Användarflöde 3 - exempel

I användarflöde 3 ska ett existerande eller framtida skick till ett nationellt register skapas.

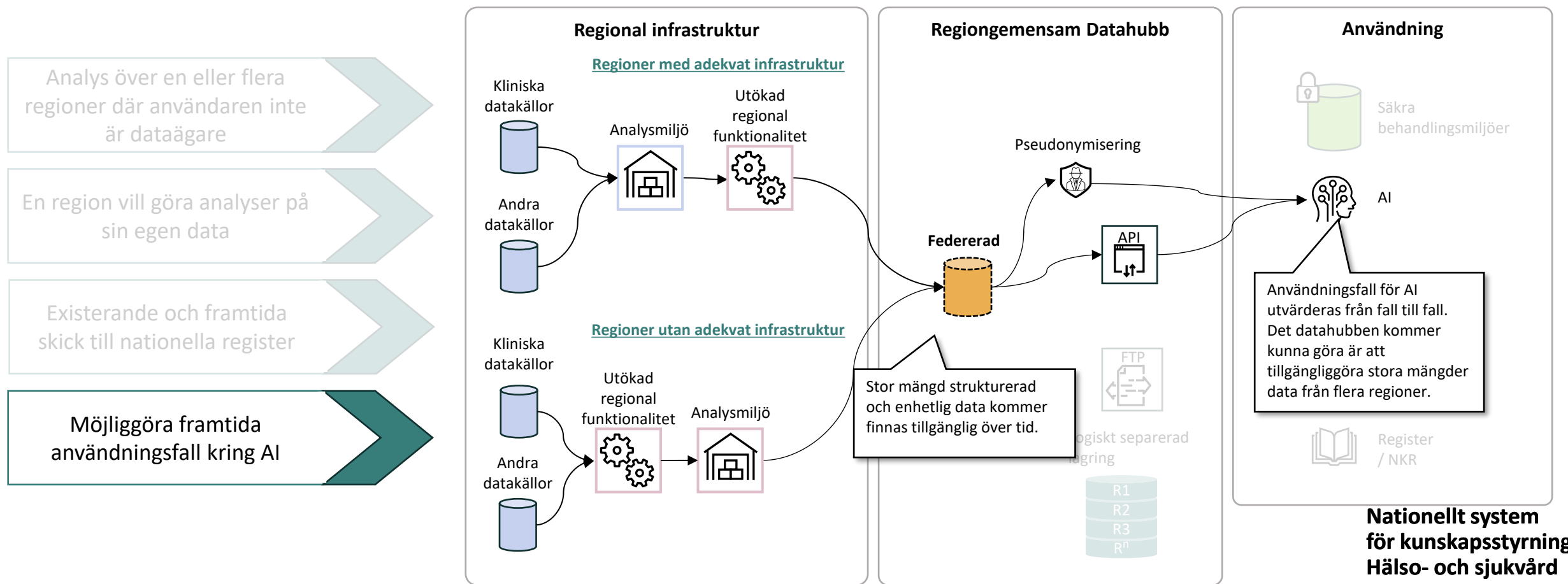
- Befintlig regional funktionalitet
- Utökad regional funktionalitet
- Regiongemensam funktionalitet
- Användning



Användarflöde 4 - exempel

Användarflöde 4 handlar om att möjliggöra AI. Referensarkitekturen är inte specifikt designad för att realisera AI-användningsfall men stora mängder data kommer över tid finns tillgänglig via hubben vilket möjliggör AI.

- Befintlig regional funktionalitet
- Utökad regional funktionalitet
- Regiongemensam funktionalitet
- Användning



**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Dataintegrering

Dataintegrering

Om dataintegrering

Dataintegrering är processen att samla in, harmonisera och konsolidera data från olika källor till en enhetlig och tillgänglig informationsmodell. Den omfattar metoder för att samla in och kombinera data, oavsett om källorna är interna eller externa, lokala eller molnbaserade, för att möjliggöra rapportering, analys och operativa processer.

Dataintegrering skapar en sammanhängande datavy som underlättar interaktion för användare och system.

Utvalda metoder

1. **Dataintagning:** Processen där data extraheras från källsystem och fysiskt kopieras till en central lagringsmiljö för vidare bearbetning eller analys.
2. **Datafederering:** Processen att hämta och temporärt kopiera data från olika källsystem till en central bearbetningsmiljö vid en specifik förfrågan. Genom tekniker som API:er eller distribuerade frågemotorer kombineras data dynamiskt för att skapa en enhetlig vy för analys, rapportering eller för att användas i operativa processer. Data lagras inte permanent i den centrala miljön, vilket möjliggör flexibel och behovsdriven dataintegration.
3. **Datavirtualisering:** Skapar en abstrakt, enhetlig vy över data från flera källsystem utan att flytta data. Ett virtuellt lager används för att i realtid hämta och presentera data som en sammanhängande resurs, oavsett om källorna är lokala, molnbaserade eller heterogena. Detta underlättar snabb och skalbar åtkomst för användare och system, utan behov av permanent datalagring eller fysisk konsolidering.

Dataintegrering i en regional kontext

För att minimera påverkan på regionernas aktiva arbeten kring data kommer regionernas befintliga infrastrukturer användas. Detta innebär att existerande dataintegrationer mellan källor (vårdsystem och andra) och regionernas analysverktyg kommer användas i så stor utsträckning som möjligt. Istället kommer primärt data-integration mellan regionernas analysverktyg och den regionsgemensamma datahubben att eftersträvas. Regionernas dataanalysmiljöer blir således gränssnittet mot datahubben.

Dataintegrering i den regiongemensamma datahubben

Oavsett vilken metod som väljs kommer dataintegrering i någon form att vara nödvändig för att uppnå syftet med en regiongemensam datahubb. Data från regionerna måste delas(ex. genom SFTP) och göras tillgänglig(ex. genom API) för att möjliggöra hubbens funktion.

I de följande avsnitten analyseras de utvalda koncepten utifrån ett regionalt perspektiv. För att hålla fokus på den centrala frågan – om data ska lagras permanent i den regiongemensamma datahubben eller inte – **kommer datafederering och datavirtualisering att behandlas gemensamt i den kommande analysen.**

Regiongemensam datahubb – Federerad eller central lagring

Fördjupning om Federering

Inledning

Federering innebär att data integreras från flera typer av källor utan att permanent flyttas till en central lagringspunkt. Det federerade systemet skapar en enad vy från alla datakällor genom ett virtuellt datalager ovanpå och att data hämtas och kombineras från olika källor när den ska observeras. Med andra ord skapas antingen ett virtuellt datalager exempelvis genom metadata eller så hämtas data och lagras temporärt när den ska observeras.

Utvalda koncept

1. **Virtuell integration:** Data lagras enbart hos källan vilket undviker problem med att duplicera och bibehålla data på flera platser samt minimerar lagringskostnader.
2. **Flexibel dataintegration:** Datafederering kan hantera olika typer av källor (relationsdatabaser, NoSQL system, API:er etc.).
3. **Enhetligt användargränssnitt:** Tillåter användaren att hämta data genom ett enhetligt språk, vanligtvis SQL, GraphQL eller något liknande.

Vanliga användningsfall för datafederering

- **Många avdelningar eller dotterbolag:** Genom federering är det möjligt att ha realtidsrapportering från olika avdelningar medan både faktisk data och ägandeskapet för den data ligger kvar hos respektive enhet.
- **Hälsa- och forskning:** Möjliggöra analys av känslig data utan att data behöver flyttas och säkerställer att patientdata fortsätter kontrolleras av dataägaren.
- **Hybrid Cloud och On-prem system:** Kan undvika komplexa överföringar mellan olika miljöer. Anrop kan skrivas en gång och användas flera gånger på olika källor. Kan vara ett effektivt verktyg och mellansteg vid en cloud-migrering.

Varför ska federering användas i en regiongemensam miljö?

Datafederering möjliggör åtkomst till regionernas data utan att lagra dem permanent i en central datahubb. Tre huvudsakliga fördelar har identifierats:

Förenklad hantering av datakvalitet

Genom att data förblir inom regional infrastruktur undviks behovet av att bearbeta datakvalitet i en central lagringsmiljö. Detta minskar risken för inkonsekvenser och fokuserar kvalitetsarbetet på regionernas egna system.

Bevarad regional autonomi och kontroll

Eftersom data inte flyttas till en central miljö behåller regionerna fullt ägandeskap och kontroll över sina data. De kan själva styra åtkomst, delning och datastyrning enligt regionala policyer.

Enhetlig åtkomst via gemensamt användargränssnitt

Datafederering skapar ett gemensamt gränssnitt som gör regionernas heterogena datamiljöer tillgängliga som en enhetlig resurs. Detta förenklar för användare att arbeta med data utan att behöva hantera tekniska skillnader mellan regionerna.

Exempel: Ineras infrastruktur för informationsdelning

Om Ineras Nationella tjänsteplattform

- Ineras nationella tjänsteplattform är en knutpunkt för utbyte av data inom vård och omsorg. I dagsläget hanterar Ineras nationella tjänsteplattform 600 miljoner anrop per månad och 1500 vårdgivare tillgängliggör patientdata via den.
- Den data som tillgängliggörs följer Ineras nationella tjänstekontrakt som stipulerar innehåll och format.
- Det finns över 200 tjänstekontrakt, varav drygt 20 bär patientdata inom området journal och läkemedel. Dessa tjänstekontrakt utgör tillsammans med den nationella tjänsteplattformen det fundament som e-tjänsterna Nationell patientöversikt, Journalen i 1177 och Informationsförsörjning av nationella kvalitetsregister vilar på.
- Förutom dessa 20 rena JoL-tjänstekontrakt finns det betydligt fler tjänstekontrakt som bär på patientdata och annan data som är intressant i sekundäranvändningssyfte.

Ineras Nationella tjänsteplattform som datakälla

Informationsutbytet över den nationella tjänsteplattformen följer Ineras samverkansarkitektur som beskriver hur informationsutbyte ska ske inom vård- och omsorgssektorn.

Samverkansarkitekturen förordar en federerad modell där informationen inte lagras centralt utan läses från respektive region/vårdgivare/kommun när den ska visas upp i NPÖ eller lämnas ut till ett nationellt kvalitetsregister (via NKRR). Denna arkitektur ligger i linje med den federerade arkitektur som föreslås för vårddatahubben.

Eftersom den nationella tjänsteplattformen redan idag används för utlämnande av information från vårdgivare till kvalitetsregister ligger det nära till hands att se den som en potentiell datakälla för sekundäranvändning i ett bredare perspektiv. Att bredda användningen till hela det behov som finns inom sekundäranvändning bär dock på vissa utmaningar.

Ineras nationella tjänsteplattform och tjänstekontrakt som datakälla för bred sekundäranvändning

Utmaningar	<p>Täckningsgraden kunde vara bättre både på bredd (antalet informationsmängder som täcks in av tjänstekontrakten) och djup (antalet parter som tillgängliggör sin information via tjänstekontrakten).</p> <p>Tjänstekontrakten är utformade för att bära information om en (1) patient, vilket fungerar bra för att informationsförsörja nationella kvalitetsregister och även andra användningsfall. Det är dock inte ändamålsenligt för all sekundäranvändning. Detta är relativt enkelt att åtgärda men kräver åtgärder i varje källsystem, vilket kräver beställningar till systemleverantörer.</p> <p>Vissa källsystem som idag tillgängliggör information kan få utmaningar med prestanda om man ska kunna läsa många patienters data i varje anrop.</p>
Möjligheter	<p>Den är väletablerad, inte minst för nationella kvalitetsregister.</p> <p>Den bygger på rätt koncept; federation.</p> <p>Täckningsgraden kunde vara bättre men informationsdelningen omfattar redan mycket så pass stora datamängder att det är en värdefull datakälla för sekundäranvändning.</p> <p>Förhållandevis enkelt att utöka både antalet tjänstekontrakt och anslutna parter.</p>

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Vad är central lagring?

Om central lagring

Central lagring innebär att data från många olika källor samlas in och lagras i en gemensam central miljö. Det skulle kunna vara en databas, ett data warehouse, en data lake eller en liknande centralt system. Nyckeln är att data fysiskt kopieras från sina källsystem till en central miljö och är därför inte längre är bundet till en källa. Framtida behov av data kan nu tillgodoses via en central miljö.

Utvalda koncept

1. **ETL/ELT processer:** Är ett arbetsflöde som handlar om att Extrahera, Transformera och Lagra data.
2. **Batch processing:** För det mesta uppdateras data efter ett regelbundet schema, ibland så ofta som varje minut, men sällan inte i realtid.
3. **Centraliserad säkerhet och compliance:** Data i en central miljö möjliggör standardiserad hantering av åtkomstkontroller, kryptering och loggning. Detta effektiviserar säkerhetsarbetet och underlättar efterlevnad av regelverk genom förbättrad spårbarhet och enhetlig datatillgång.

Vanliga användningsfall för central lagring

- **BI-rapportering:** Data från flera system konsolideras för att skapa rapporter och analyser som stöder uppföljning av verksamhetens prestationer. Konsolideringen möjliggör enhetlig datatillgång och underlättar datadrivna beslut.
- **Behålla historisk data:** Genom att exportera data till en central lagringsmiljö kan stora mängder historisk data bevaras över tid, vilket ofta inte är möjligt i operationella system.
- **Maskininlärning och AI:** Genom att konsolidera stora datamängder i en enhetlig lagringsmiljö möjliggörs avancerade tillämpningar av maskininlärning och AI. Dessa teknologier drar nytta av omfattande data för att träna modeller och generera insikter.

Varför ska central lagring användas i en regiongemensam miljö?

Tre huvudsakliga anledningar har identifierats:

Beprövad metod

Tekniken har använts över många år och det finns en stor mängd leverantörer och open-source verktyg tillgängliga på marknaden.

Hög analysprestanda

Eftersom data lagras centralt kan många optimeringsmetoder, exempelvis indexering och caching, användas för att säkerställa höga prestanda.

Enhetlig säkerhet och compliance

Det är enklare att underhålla en enhetlig säkerhetsmodell för åtkomsthantering, loggar, pseudonymisering etc. istället för att koordinera detta över flera system.

Jämförelse central och federerad lagring

Område	Central	Federerad
Grundkoncept	Kopiera data till en central lagring	Data stannar hos källan
Tillgänglighet	Beror på uppdateringsfrekvens; kan innebära fördröjningar i datatillgänglighet	Nära realtid
Skalbarhet	Beror på val av system och leverantör men generellt god	Beror starkt på underliggande system
Flexibilitet	Mindre flexibel då data som ska användas måste laddas in	Väldigt flexibel och goda möjligheter att hantera ad-hoc förfrågningar
Kontroll	Lokal organisation har begränsad operativ kontroll, vilket kräver tydliga avtal för att reglera åtkomst och användning.	Lokal organisation behåller full operativ kontroll över sin data och kan styra åtkomst direkt.
Datakvalitet	Data dupliceras vilket ställer höga krav på centrala kvalitetsprocesser utöver de regionala	Data ligger kvar ute i källorna vilket ställer höga krav på regionala kvalitetsprocesser
Metadata	Enklare genom central metadatahantering i en central katalog, men kräver inläsning och harmonisering av regionernas metadata vid konsolidering	Komplex – metadata behöver federeras dynamiskt från olika källor
Datakonsistens och synkronisering	Risk för inkonsistens mellan lokala källor och central lagring om data ändras lokalt efter kopiering; kräver synkroniseringsprocesser	Data är alltid aktuell eftersom den hämtas direkt från källorna; ingen risk för synkroniseringsproblem

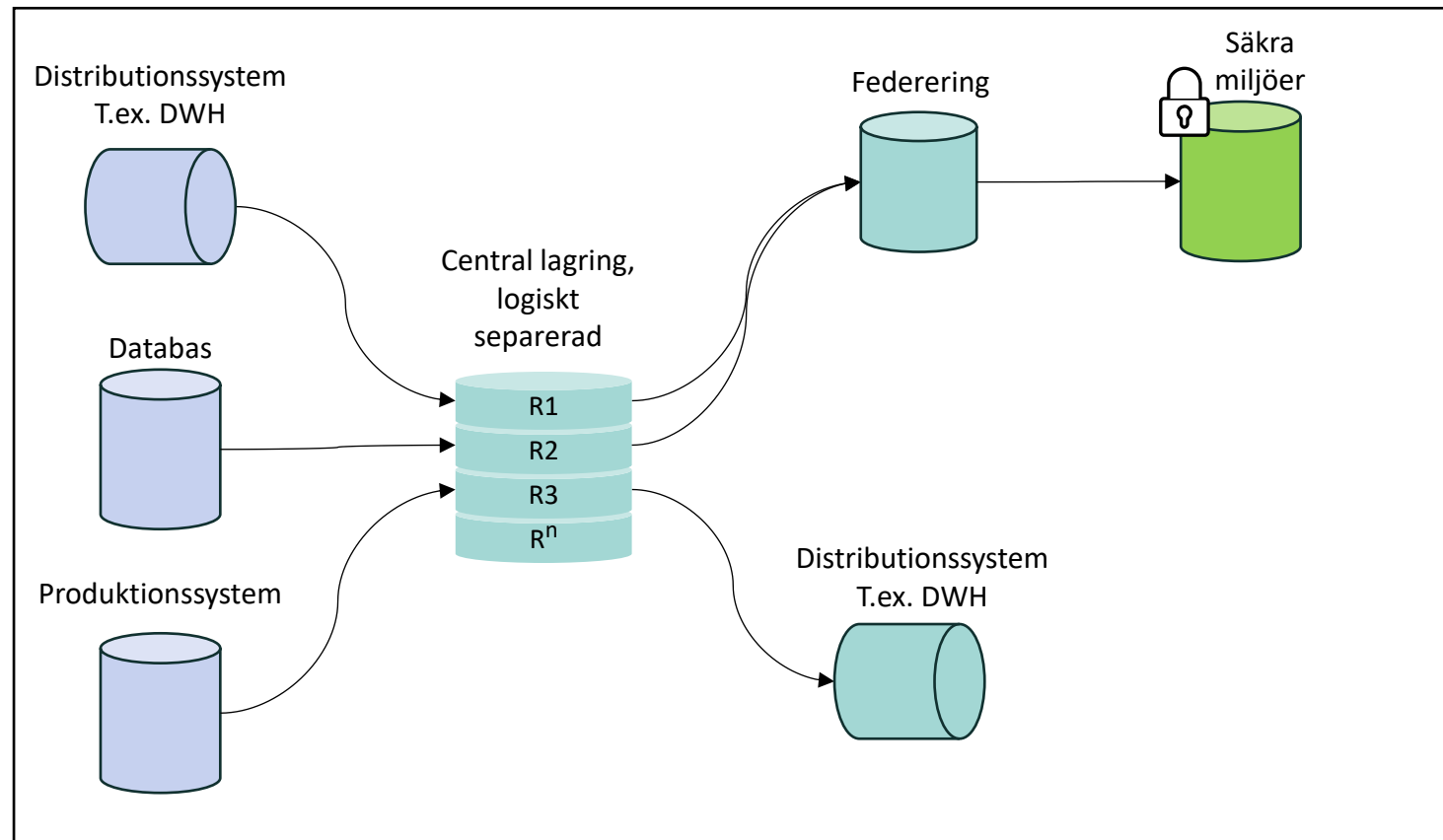
Slutsats

Båda lagringskoncepten har användningsfall där de är lämpliga, med styrkor som kompletterar varandra. I den regiongemensamma datahubben bedöms en kombinerad lösning vara mest lämpad, där viss data lagras permanent centralt medan annan hanteras federerat. Anledningen är datahubbens komplexa krav, med olika användningsfall som BI-rapportering och maskininlärning, samt heterogena källor såsom databaser och äldre system. En kombinerad lösning kan balansera prestanda och regelefterlevnad från central lagring med flexibilitet, autonomi och lokalt ansvarstagande från federering. I nästa segment jämförs fyra alternativ för hur central och federerad lagring kan kombineras för att möta dessa behov.

**Nationellt system
för kunskapsstyrning
Hälso- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Alternativ 1: Central lagring



Beskrivning

Använder traditionell ETL/ELT flöde för att flytta data till en regiongemensam lagringsyta. Lagringen är logiskt separerad så varje regions data är isolerad. De federerade frågorna används för att sammanställa data från flera regioner vid behov om lagliga förutsättningar finns



Fördelar

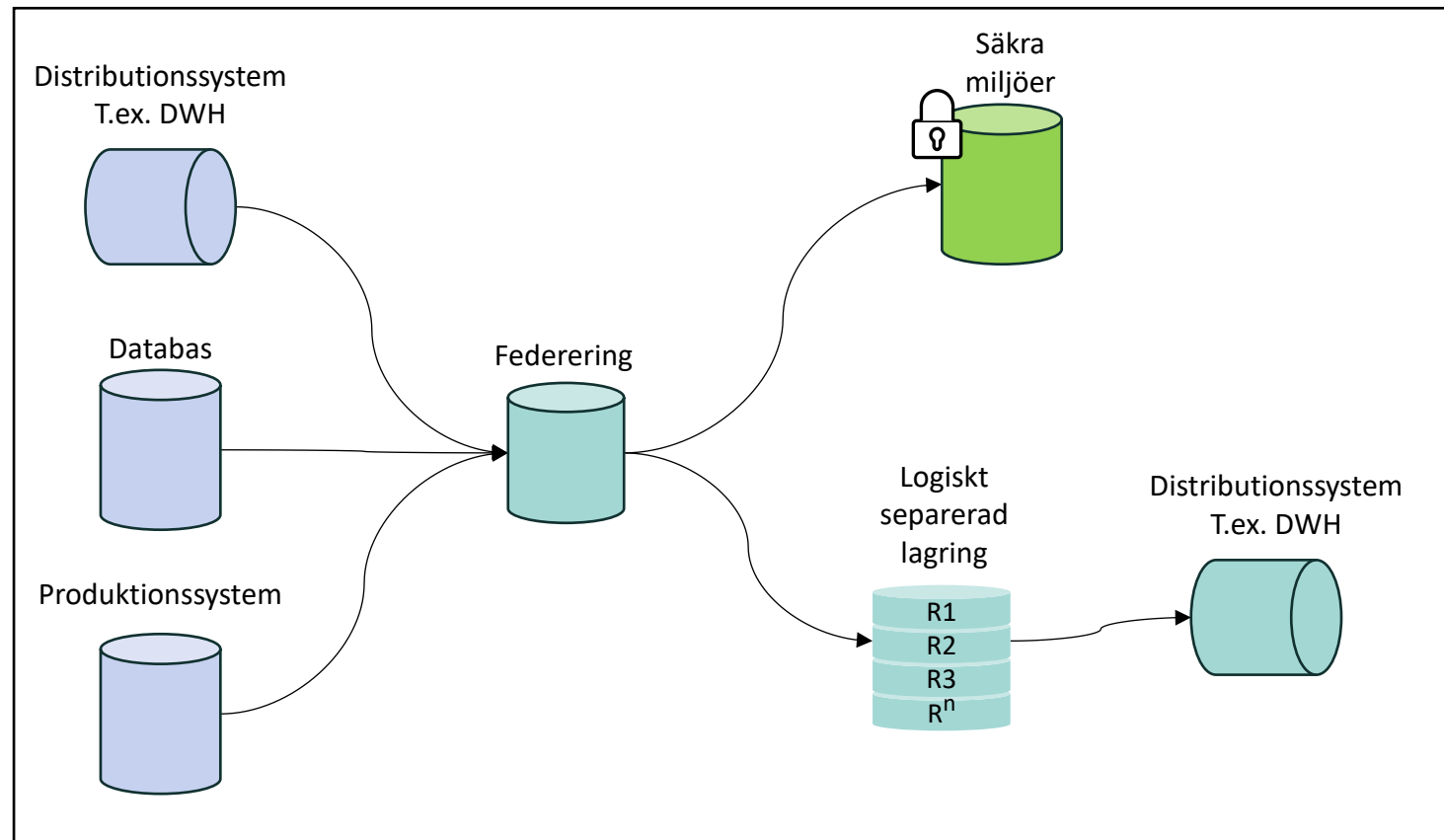
- Beprövat upplägg med liknande initiativ att ta inspiration av (t.ex. GMS)
- Goda möjligheter att analysera och skapa t.ex. rapporter
- Laddningar kan anpassas efter källsystem för att minimera last på system



Nackdelar

- Kan bara analysera data som lagrats
- Data laddas i batcher så tillgänglighet är beroende på frekvens
- Data är inte alltid aktuell
- Data lagras både hos regionerna och regiongemensamt vilket innebär att arbete krävs för att säkerställa kvalitet

Alternativ 2: Federering



Fördelar

- Ingen duplicerad lagring av data
- Kontroll av data stannar hos källan
- Live förfrågning på operationell data
- Metadatadrivet
- Kan hantera stora mängder data genom ex. caching-lösningar



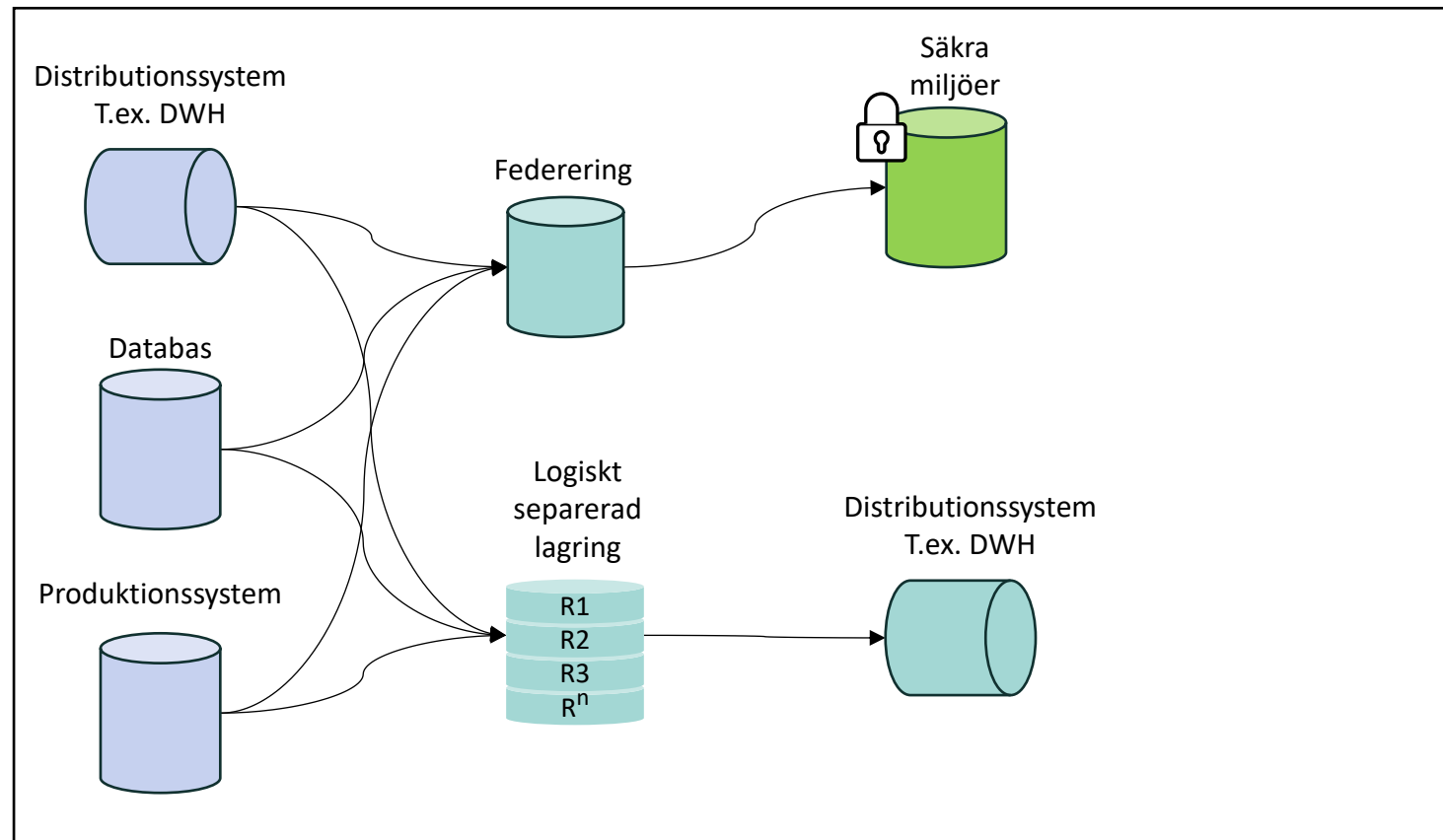
Nackdelar

- Datavirtualiseringen är inte gjord för att hantera stora batchprocesser
- Svarstider kan bli ett problem
- Beroende av fungerande kommunikation till regional infrastruktur
- Om förfrågningar är sub-optimala kan det påverka prestanda för produktionssystem

Beskrivning

Federering används som den primära kopplingen mot regionala miljöer och från den federerade miljön distribueras datamängder beroende på användningsfall. Om ett användningsfall behöver det kan central lagring användas.

Alternativ 3: Parallellt



Beskrivning

En kombination av båda tidigare beskrivna upplägg. Varje system har en koppling till både en federerad miljö och en central lagring.



Fördelar

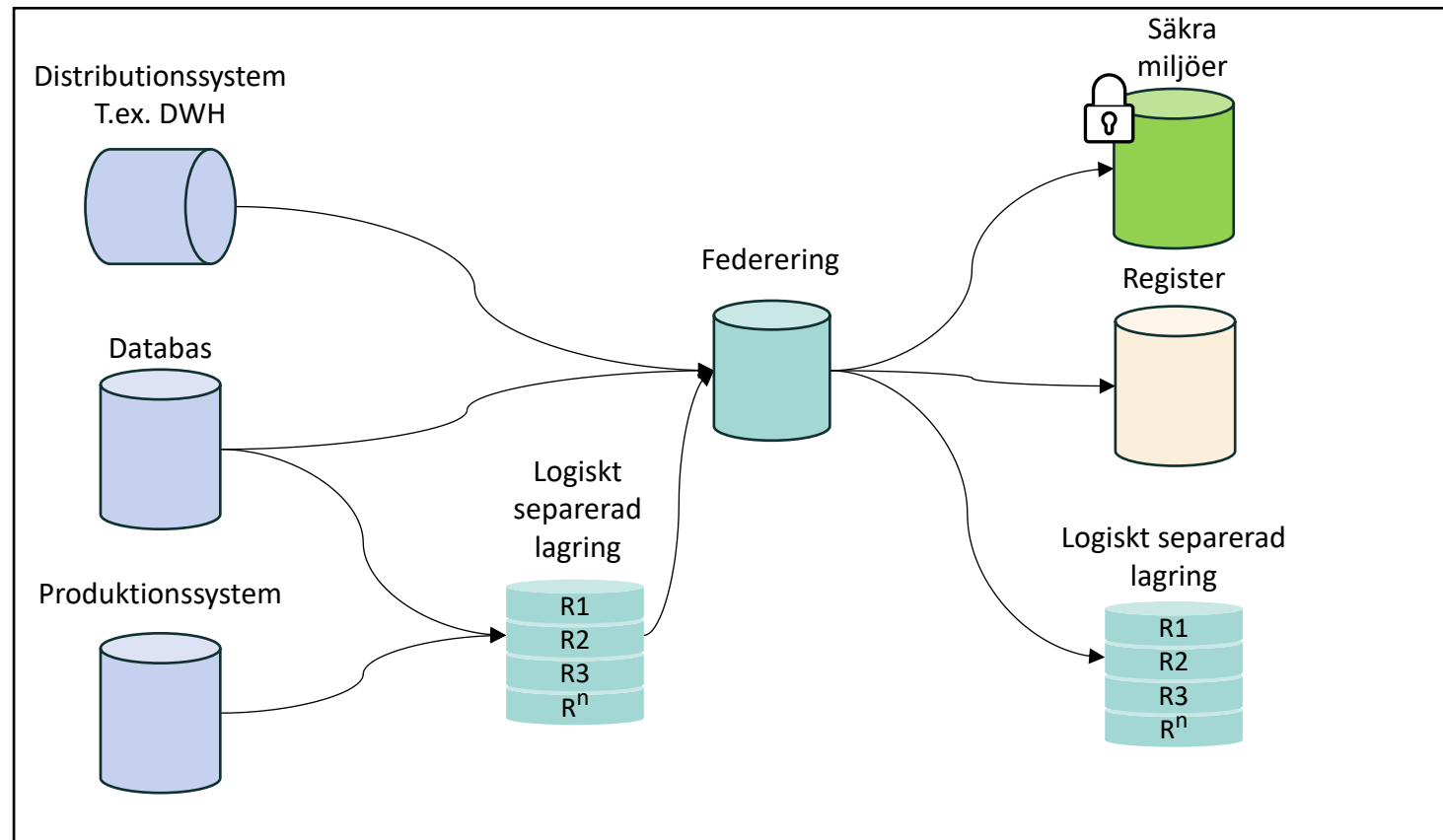
- Kan hantera både realtid och stora batchladdningar
- Kan hantera strukturerad och ostrukturerad data
- Stor flexibilitet och förmåga att hantera alla typer av användningsfall



Nackdelar

- Två separata system behöver utredas, implementeras och underhållas
- Belastning på källsystem på grund av dubbla systemkopplingar
- Dubbla infrastrukturer av systemkopplingar att underhålla
- Data lagras både hos regionerna och regiongemensamt vilket innebär arbete krävs för att säkerställa kvalitet

Alternativ 4: Hybrid



Fördelar

- Nyttjar befintlig infrastruktur inom regionerna
- Duplicerar inte data i onödan
- Belastar inte produktionssystem i onödan
- Kan hantera strukturerad och ostrukturerad data
- En "koppling" per system
- Goda möjligheter att modellera och transformera data
- Ställer lägre krav på part som vill ansluta



Nackdelar

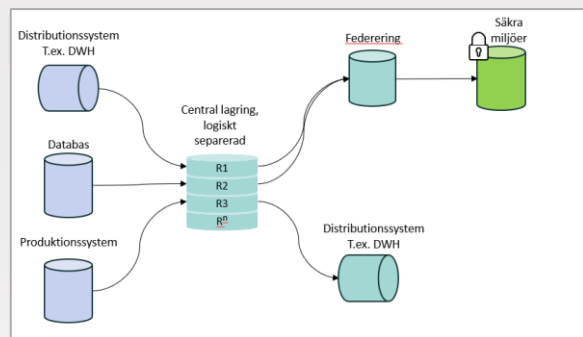
- Noggrann undersökning av leverantörer för att säkerställa att alla system täcks
- Komplexare lösning jämfört med att enbart använda fysisk lagring eller federering

Beskrivning

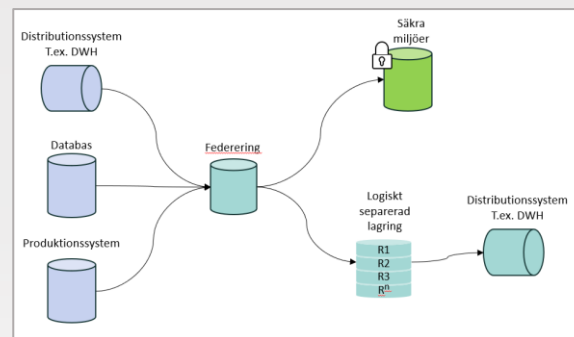
En kombinerad lösning där källas typ avgör användningen av ett federerat flöde eller ett batch-flöde. Distributionssystem och icke-kritiska databaser använder ett federerat flöde medan kritiska databaser och produktionssystem använder ett batchflöde till fysisk lagring som i sin tur federeras.

Projektet har utvärderat fyra olika arkitekturella lösningar för den regiongemensamma datahubben

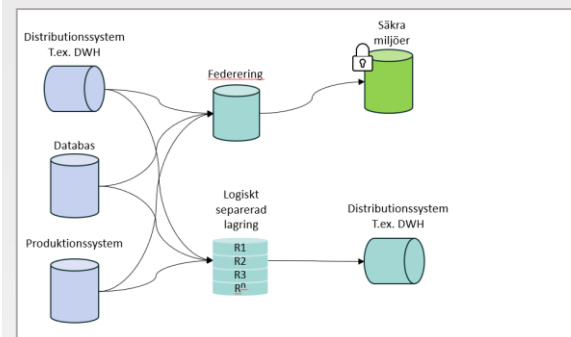
Central lagring



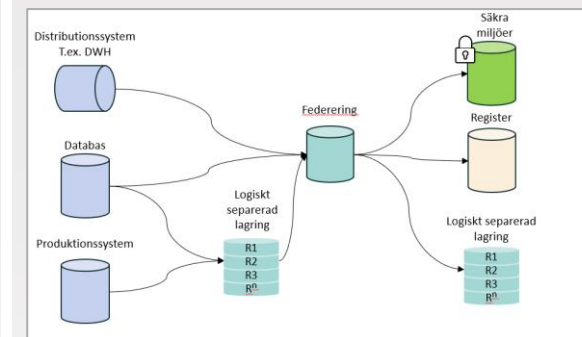
Federering



Parallellt



Hybrid



Central lagring lämpar sig för en del användningsfall och datamängder. För datahubben lämpar det sig mindre bra eftersom kontroll av data flyttar från källsystem till den centrala lagringsplattform. Datahubben strävar efter att regionerna ska bibehålla kontroll i så stor utsträckning som möjligt vilket gör att detta alternativ inte rekommenderas

Federering lämpar sig väl som lösning för datahubben. Genom federering kan regionernas existerande infrastrukturer, många med distributionssystem, nyttjas. Central lagring kan sedan användas efter den federerade för specifika datamängder där en batchhämtning är mer lämplig. Denna lösning rekommenderas.

En parallell lösning innebär flera kopplingar från regionernas system till datahubben vilket kommer innebära mer arbete för regionerna att implementera och underhålla. Utöver detta kommer det vara två stycken system som kräver resurser vilket gör att detta alternativ inte rekommenderas

En hybrid lösning ger många av de fördelar som en federerad lösning har samtidigt som många av de potentiella problemen minimeras. Däremot skulle vissa datamängder flyttas till central lagring vilket skulle göra att kontrollen av data flyttas. **Andra möjligheter för att återskapa samma funktionalitet regionalt bör undersökas innan denna typ av lösning rekommenderas.**

Resonemang

Slutsats



Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård

SVERIGES REGIONER I SAMVERKAN

Datamodellering

Vad är datamodellering?

Om datamodellering

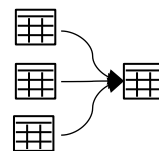
Datamodellering handlar om att skapa en beskrivning av datastrukturer och deras relationer inom en miljö. Det kan liknas med en ritning som visar hur data ska lagras, struktureras och manipuleras i databaser och/eller data warehouses.

Utvalda koncept

- Entiteter och attribut:** Entiteter representerar objekt (t.ex. patient, vårdbesök) och attribut beskriver entiteternas egenskaper (t.ex. längd, vikt)
- Relationer:** Beskriver hur entiteter relaterar till varandra. Relationer kan vara ett-till-ett, en-till-många eller många-till-många. Ex. En patient kan ha flera vårdbesök vilket ger en en-till-många relation
- Typer av modeller:**
 - Konceptuell modell:** Överblick som fokuserar på övergripande struktur och relationer utan att innehålla tekniska detaljer
 - Logisk modell:** Mer detaljerad än den konceptuella modellen. Den logiska modellen definierar strukturen på dataelementen och relationerna men är fortfarande oberoende av specifika databashanteringssystem.
 - Fysisk modell:** Översätter den logiska modellen till en fysisk struktur som kan implementeras på ett specifikt databashanteringssystem inklusive exempelvis datatyper och index

När ett databashanteringssystem implementeras kan olika metoder användas för att modellera data. Nedan presenteras tre olika alternativ för modellering:

Gemensam

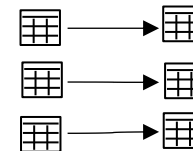


Syfte: Skapar ett standardiserat ramverk för att representera data från olika källor

Fokus: Säkerställer jämförbarhet i datastrukturer, definitioner och relationer. Vilket gör det lättare att integrera och analysera data

Användning: Passar organisationer som vill skapa enhetlighet från olika källor för analys och rapporter

Heterogen

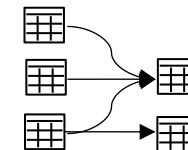


Syfte: Tillåter användning av flera datamodeller anpassade efter varje datakällas behov.

Fokus: Behåller unika attribut och strukturer från individuella källor vilket skapar en stor flexibilitet kring hur data kan användas

Användning: Passar organisationer med spridda datakällor där specifika krav måste mötas

Hybrid



Syfte: Kombinerar element från gemensamma och heterogena modeller

Fokus: Använd ett gemensamt ramverk för kärndataelement och samtidigt tillåta ytterligare attribut som är källspecifika

Användning: Balanserar behovet av enhetlig data och integration med flexibilitet för att behålla unika dataelement

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Varför är datamodellering viktigt för datahubben?

Marknadstrender

Över de senaste åren har trenden för data inom hälsosektorn varit att gå mot öppna system. Framtidens vårdinformationsmiljöer separerar ut datalagring från applikation och logik för att samma data ska vara systemoberoende och kunna användas på flera sätt. Detta är sant för såväl primär- och sekundäranvändning. För att data ska kunna användas på det här sättet krävs en strukturerad datalagring samt teknisk och semantisk interoperabilitet.

Hur ser det ut i regionerna idag?

Regionerna har över många års tid byggt upp sina infrastrukturer och som en del av dessa även skapat sina egna datamodeller. Detta resulterar i att det finns en stor spridning kring hur information lagras vilket gör att data från olika regioner idag inte går att jämföra på ett enkelt sätt.

Vilket arbete krävs för att gå från olika modeller till en gemensam?

Ett manuellt modelleringsarbete kommer behöva genomföras där olika datamängder från regionerna kartläggs och översätts till den regiongemensamma datamodellen.

Slutsats

- För att möjliggöra fler analyser och jämförelser mellan regioner behöver data vara modellerad i en **gemensam** eller **hybrid** datamodell där en hybrid modell enbart skulle appliceras om en region har unika datamängder.
- Arbetet att mappa från regionernas egna modeller till den gemensamma modellen bör ske **hos regionerna** med stöd från kompetenscentrum för att regionerna ska behålla kontroll över data samt att modelleringsarbetet ska kunna underhållas över tid.

Hur ska en gemensam datamodell skapas?

En av de största frågorna som behöver undersökas i en framtida utredning är hur en gemensam datamodell ska se ut. Tre potentiella scenarier har identifierats.

$$\frac{x}{y}$$

Utgå från **existerande regionala datamodeller** och **hitta minsta gemensamma modell**



Utgå från **internationella standarder** och modeller och applicera och anpassa dessa till en svensk kontext



Utgå från **existerande regionala datamodeller** och välj den modell som har **bredast användning**

I den fortsatta utredningen kring en gemensam datamodell bör många olika aktörer involveras, exempelvis representanter från regionerna och socialstyrelsen.

En första jämförelse mellan primär- och sekundäranvändningen kan ses i appendix* och berör bland annat datamodellering.

I kommande bild redogörs för vanligt förekommande internationella standarder som bör inkluderas i en framtida utredning kring datamodell

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Se Appendix 4

Vanligt förekommande standarder

OpenEHR

- En uppsättning standardiserade öppna **specifikationer** för lagring, hantering, hämtning och utbyte av hälso- och sjukvårdsdata. Specifikationerna används för att skapa datamodeller, kliniska riktlinjer och olika programvaror och tjänster.
- Ett särskilt fokus på långsiktig hantering och **beständighet** av elektroniska hälsojournaler(EHR-data). OpenEHR-baserade data lagras med hjälp av en semantiskt sammanhängande, öppen, internationell datamodell utvecklad av kliniskt yrkesverksamma med styrning av openEHR International.
- OpenEHR använder en modellering med två nivåer (arketyper och mallar) för att separera klinisk information från tekniska implementationer.
- Arketyper representerar kliniska och sociala vårdkoncept i ett rikt, detaljerat format. De inkluderar också kontext, till exempel om blodtrycket mättes i liggande eller sittande position.
- Utformad för att vara leverantörs-, teknik- och språkneutral.
- Möjliggör separation av applikation och data: data lagras i ett openEHR kliniskt datalager (CDR) och innehållsbaserade frågor av hög detaljeringsgrad möjliggörs.

HL7 FHIR

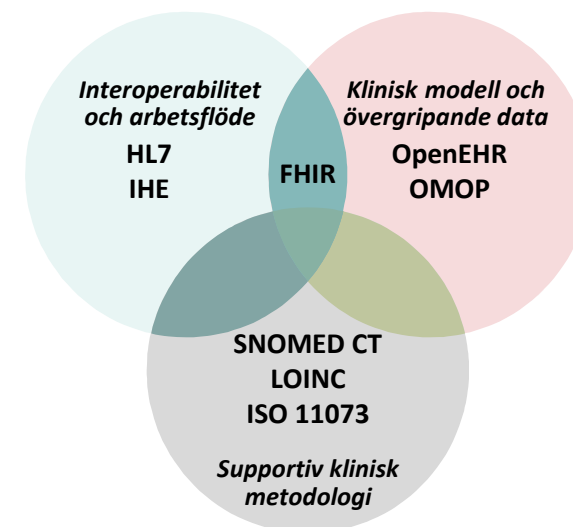
- HL7 FHIR möjliggör **datautbyte** mellan IT-system med olika interna datamodeller.
- Utformad för att göra datautbytet snabbare och mer flexibelt.
- Använder en modern webbaserad metod med RESTful API:er, fokuserar på interoperabilitet och enkel implementering.
- Inte avsedd för långsiktig datalagring, den täcker de vanligaste dataelementen för datautbyte.
- De huvudsakliga koncepten inkluderar **FHIR-resurser** och **profiler**.
 - Resurs är en grundläggande enhet av data som representerar en vårdrelaterad enhet eller objekt (t.ex. Patient, Observation, etc.).
 - Profil är en specifikation eller anpassning av en FHIR-resurs, som ofta tillämpar ytterligare regler, begränsningar eller tillägg för specifika användningsfall, regioner eller affärsbehov.

OMOP CDM

- Ett standardiserat vokabulär utformat för observationsforskning och storskalig analytisk användning av hälso- och sjukvårdsdata.
- Utformat för att möjliggöra storskalig epidemiologisk forskning och generering av verklighetsbaserad evidens, särskilt för att bedöma medicinska resultat och behandlingar.
- Standardiserar olika dataformat för konsekvens i forskningen, vilket stödjer samarbetsinriktade, flersidiga studier med en gemensam datastruktur.
- Använder standardterminologier (t.ex. SNOMED, ICD, LOINC).

Kompletterande användningsområden

- openEHR utmärker sig i att fånga detaljerad klinisk data.
- FHIR underlättar effektivt datautbyte.
- OMOP CDM möjliggör robust dataanalys och forskning.
- Dessa standarder kan användas tillsammans med varandra.



**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Hur skulle openEHR kunna appliceras för datahubben?

För vilken data lämpar sig openEHR?

- Ur källdatasynpunkt passar openEHR för beständighet av kliniska och socialvårdsrelaterade data. openEHR kan också omfatta demografiska data, men vanligtvis är demografiska servrar/system separata och de har ett HL7 FHIR-gränssnitt.
- openEHR är som bäst när det gäller att lagra och hantera kliniska data för primäranvändning. Emellertid används openEHR också för sekundäranvändning om detaljerad, rik data krävs.
- Regionala öppna plattformar inkluderar ett kliniskt datarepositorium (CDR). Data lagras där i ett arkotypiskt format. openEHR-data är pseudonymiserade genom design. Detta innebär att en patient har ett unikt EHR-ID och motsvarande demografiska data lagras separat och efterfrågas vid behov, till exempel i användargränssnittet.
- openEHR kan rekommenderas för nya regionala CDR:er för att samla in och harmonisera data från olika lokala källsystem. Dessa CDR:er skulle möjliggöra innehållsbaserade, finkorniga sökningar och insamling av data till den nationella (regiongemensamma) plattformen.
- På regional nivå är det också möjligt att kombinera CDR med en data lake för datainsamling, rensning och harmonisering. Datainsamling från äldre system varierar mycket, inklusive olika sätt att exportera data, och en data lake kan fungera som en användbar behållare före CDR:en.
- Om det finns en regional plattform tillgänglig med en proprietär datamodell, krävs en mappning till openEHR på regional nivå, alternativt inom den nationella plattformen. Det är enklast att skapa förfrågningar efter specifika data med AQL-verktyg från CDR.

Vilka övriga datamodeller kan komplettera openEHR för att täcka de flesta data som krävs för sekundäranvändning?

Klinisk data

- openEHR är grunden för klinisk pseudonymiserad data.
- HL7 FHIR kan också användas för klinisk data, men särskilt för utbyte av mer avgränsade datamängder.
- OMOP CDM inkluderar vokabulär för forskning och storskalig analytisk användning. I alla fall behövs relevanta kartläggningar.
- Om vi använder openEHR som den harmoniserade datamodellen kan vi minska de kartläggningar som behövs när data konverteras till FHIR och OMOP CDM. Det kan också finnas andra format i samband med den kommande implementationen av EHDS.

Icke-klinisk data

- Icke-klinisk data avser information som inte är direkt relaterad till patientvård men som är avgörande för vårdenhetens övergripande funktion. Detta kan inkludera data relaterad till administration, ekonomi, personalresurser, drift och facilitetsförvaltning.
- Det finns många standardramverk för att täcka icke-klinisk data.
- Till exempel kan HL7-standarder som HL7 Version 2.x och HL7 FHIR stödja utbyte av icke-klinisk data, såsom fakturering och administrativa arbetsflöden. IHE-integrationsprofiler stödjer arbetsflöden som inkluderar icke-klinisk data också.

Standard för BI och rapportering

- Det finns också standarder för Business Intelligence och rapportering för att analysera icke-klinisk data (t.ex. ekonomi, personalresurser) för rapporterings- och beslutsfattande syften.
- OLAP (Online Analytical Processing) är teknik för att hantera komplexa frågor med hög hastighet eller utföra multidimensionell analys på stora datamängder i ett datalager eller annan databas.
- Vårdenheter kan använda OLAP-verktyg och standarder för datalagring för att analysera icke-klinisk data.

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Datasäkerhet

Vad är datasäkerhet?

Om datasäkerhet

Datasäkerhet innefattar alla processer, teknologier, och policys som användas för att skydda digital information från obehörig åtkomst, korruption eller intrång. Konceptuellt kan det beskrivas som att data bara görs tillgänglig för rätt användare, i rätt kontext till rätt syfte.

Utvalda koncept

- **Behörighetshantering:** Varje användare eller system behöver autentiseras och ges korrekta rättigheter.
- **Kryptering:** Kryptering av data i vila med starka krypteringsstandarder, data i rörelse genom kommunikation protokoll och nyckelhantering för att hantera krypteringsnycklar
- **Säkerhetsklassning:** Åtkomst baseras på användarens attribut (ex. region, avdelning) och dataattribut (ex. säkerhetsklass, känslighet eller ägandeskap)
- **Pseudonymisering:** All tillgång övervakas och loggas för att säkerställa compliance, stötta IT-revision, och identifiera potentiell felanvändning
- **Dataintegritet:** Övervaka och kontrollera ändringar och att inga otillåtna ändringar görs av data. Spåra ursprung, transformations och transport av data. Säkerställ att loggar som innehåller informationen inte kan modifieras otillåtet.
- **Nätverkssäkerhet:** Beskriver exempelvis segmentering av nätverk för att isolera känsliga miljöer genom brandväggar med mera.

Hur ska datahubben hantera datasäkerhet?

I det första skedet är det viktigaste att genomföra en komplett datakonsekvens- och riskanalys. Analysen ska innefatta en tydligare bild kring hur informationsflödet ska fungera och en analys av var information kan attackers, innefattandes både i vila, vid förflyttning och vid användning. Den genomförda analysen bör vara ett levande arbete där nya analyser och utredningar görs kontinuerligt för att säkerställa att inte nya risker introduceras när datahubben växer.

Utöver detta bör behörighetshantering och pseudonymisering för den regiongemensamma datahubben utredas ytterligare, se separata kapitel för mer information kring det.

Gällande nätverkssäkerhet bör konceptet "never trust, always verify" undersökas vidare. Konceptet innebär att kontinuerlig autentisering och auktorisering krävs oberoende var förfrågan kommer ifrån.

Slutsats

- En datakonsekvens- och riskanalys bör genomföras i ett tidigt skede
- I uppstartsfasen under en implementation bör ett strukturerat arbete ske kring datasäkerhet för att undersöka exempelvis nätverkssäkerhet

Pseudonymisering

Vad är pseudonymisering?

Om pseudonymisering

Pseudonymisering innebär att byta ut identifierbar information (namn, personnummer, email adresser etc.) mot artificiella identifierare eller pseudonymer. Syftet är att göra det svårare att koppla samman attribut till en individ utan ytterligare information. Skillnaden mellan anonymisering och pseudonymisering är att pseudonymisering gör det möjligt att återidentifiera vid behov. Pseudonymisering ses som en lämplig skyddsåtgärd för att uppfylla olika skyldigheter ifråga om dataskydd.

Olika typer av pseudonymisering

Det finns flera olika sätt att pseudonymisera data där syfte och användning avgör vilken teknik som passar bäst. Nedan följer några vanliga exempel.

- **Token:** Innebär att personliga data byts ut mot icke-känsliga motsvarigheter (tokens). Dessa sparas på en säker plats för att kunna användas för återidentifiering.
- **Hashing (med salt):** Genom att applicera en hash-funktion skapas en sträng som inte har något koppling till den ursprungliga värdet. Genom att addera salt (ett slumpmässigt värde) kan säkerheten ökas ytterligare. Kan vara svårt att återidentifera.
- **Kryptering med nyckelhantering:** Personliga identifierare krypteras med hjälp av en nyckel. Data kan enbart dekrypteras genom tillgång till den nyckeln.
- **Datamaskering:** Data maskeras genom att delar av den döljs, exempelvis genom att enbart visa de sista siffrorna av ett telefonnummer. Detta är en mindre säker metod som har specifika användningsområden, främst kopplat till pseudonymisering för användargränssnitt.

Viktiga frågor att vara medveten om

Anses pseudonymiserad data fortfarande vara personuppgifter?

Ja, enligt GDPR anses pseudonymiserad data där individer kan återidentifieras med hjälp av ytterligare information fortfarande vara personuppgifter. Detta gäller även om den ytterligare information som krävs hålls av en tredje part

Hur hjälper pseudonymisering organisationer att hantera data under GDPR?

Pseudonymisering har flera fördelar – bland annat minimeras säkerhetsrisker genom att data inte är direkt identifierbar utan ytterligare information. Att data inte är identifierbar gör det också lättare att säkerställa att data bara används för rätt syfte.

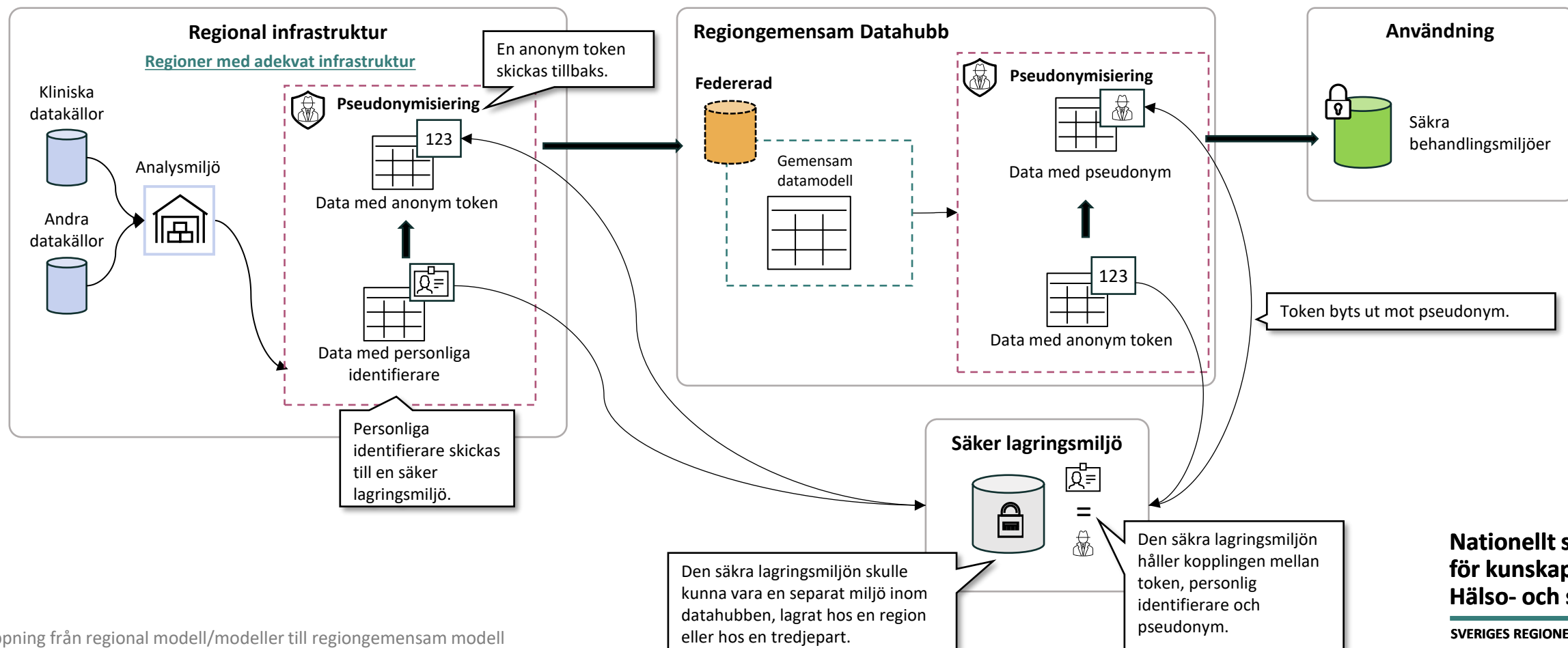
Utvalda koncept

1. **Personliga identifierare:** Inkluderar direkta identifierare som namn, telefonnummer och personnummer men även indirekta identifierare som födelsedatum, postnummer eller andra typer och kombinationer av data som kan leda till identifiering.
2. **Pseudonym:** Värdet som ersätter den personliga identifieraren.

Exempel på pseudonymisering med token

En förenklad bild av realiseringsexemplet används för att visa hur pseudonymisering kan gå till. I detta fall har det övre flödet för regioner med befintlig infrastruktur valts ut men konceptet är detsamma för regioner utan befintliga infrastruktur. Notera att syftet med denna bild är inte att vara fullständig och täcka alla scenarier utan att exemplifiera hur pseudonymisering med token kan gå till.

- Befintlig regional funktionalitet
- Utökad regional funktionalitet
- Regiongemensam funktionalitet
- Användning

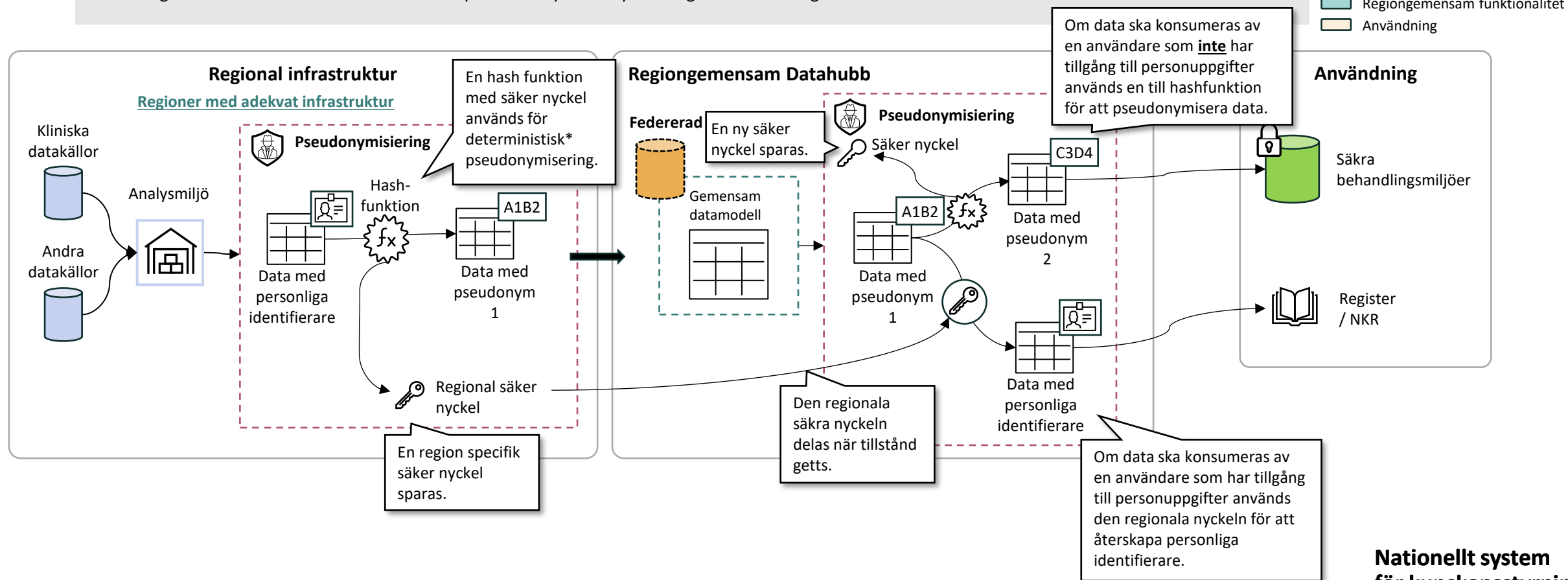


**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Exempel på pseudonymisering med hash

En förenklad bild av realiseringsexemplet används för att visa hur pseudonymisering med hash kan gå till. Notera att syftet med denna bild inte är att vara fullständig och täcka alla scenarier utan att exemplifiera hur pseudonymisering med hash kan gå till.

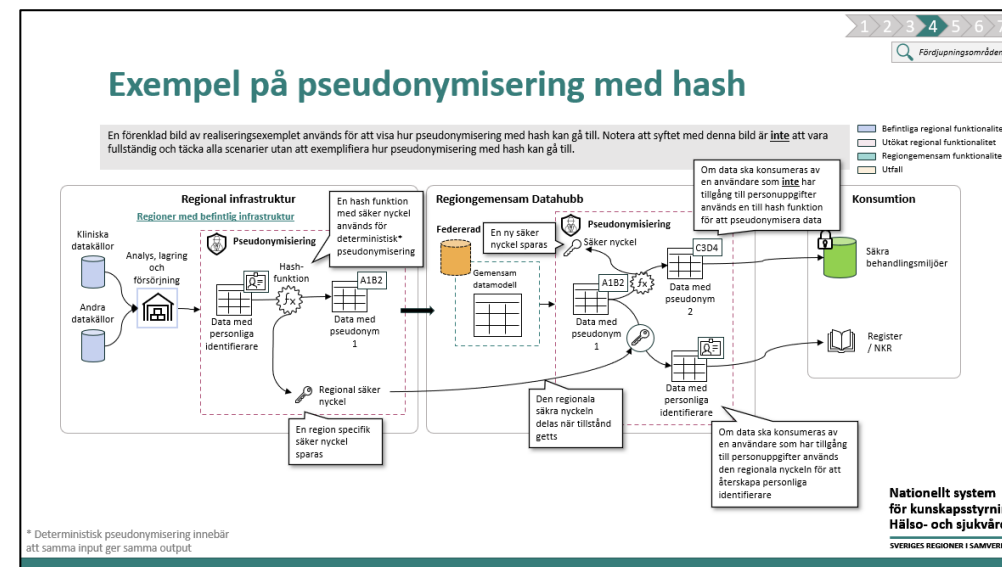
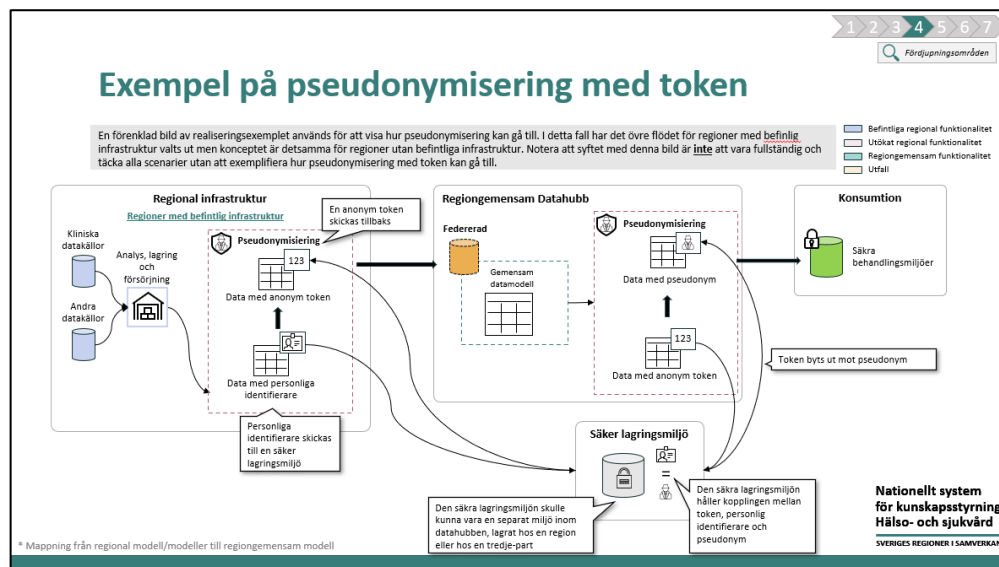


* Deterministisk pseudonymisering innebär att samma input ger samma output

Pseudonymisering

Diskussion

Syftet med exemplen i föregående bilder är att påvisa att olika typer av tekniker kan fungera i den föreslagna referensarkitekturen utan att i stor utsträckning påverka andra komponenter. Vad som kommer avgöra vilken teknik som lämpar sig bäst kommer i större utsträckning bero på vad som tillåts enligt rådande lagar.



Slutsats

- Olika pseudonymiseringstekniker kan appliceras för att uppnå tänkt syfte i den regiongemensamma datahubben.
- En juridisk utredning behöver genomföras för att tydliggöra vilka krav finns på pseudonymisering och anonymisering av data. Baserat på resultatet av utredningen kan ett beslut fattas kring vilken teknik som lämpar sig bäst i vårddatahubben.

**Nationellt system
för kunskapsstyrning
Hälso- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Behörighetslösning

Vad är en behörighetshantering? (1/2)

Om behörighetshantering

Behörighetshantering beskriver processen att kontrollera och styra vem som har tillgång till data, system eller resurser och under vilka omständigheter. Det innefattar autentisering (verifiera identiteter) och att auktorisering (ge korrekta rättigheter) för att säkerställa att enbart användare med tillstånd kan utföra specifika aktiviteter med data.

Att säkerställa effektiv och strukturerad behörighetshantering är kritiskt för datasäkerhet, regel- och policyefterlevnad (ex. GDPR) och för att garantera informationens integritet.

Utvalda koncept

- **Identitets- och åtkomsthantering (IAM):** Varje användare eller system behöver autentiseras och ges korrekta rättigheter.
- **Rollbaserad åtkomstkontroll (RBAC):** Användare blir tilldelade roller (ex. Analytiker, Data Steward eller Data scientist) och dessa roller definierar vilken typ av data och vilka aktiviteter som är tillåtna.
- **Attributbaserad åtkomstkontroll (ABAC):** Åtkomst baseras på användarens attribut (ex. region, avdelning) och dataattribut (ex. säkerhetsklass, känslighet eller ägandeskap).
- **Åtkomstloggning och monitorering:** All tillgång övervakas och loggas för att säkerställa compliance, stötta IT-revision, och identifiera potentiell felanvändning.

Hur och var är behörighetshantering viktigt i datahubben?

Regionala infrastrukturer – regionerna styr vem som har tillgång till deras interna infrastrukturer och verktyg men också vilken data som är tillgänglig i den regiongemensamma datahubben. Detta innebär att regionerna kommer ha kontroll över sin egen data och kan kontrollera vilken data som är tillgänglig när i datahubben.

Regiongemensam datahubb – i datahubben kommer behörighet bland annat hantera hur datahubben interagerar med regionala infrastrukturer och hur slutanvändare får tillgång till data. RBAC/ABAC kan användas för att säkerställa att regionerna har tillgång till sin egen data i datahubben även efter att exempelvis ytterligare modellering gjorts för att skicka data till nationella register.

Datakatalog – eftersom datakataloger innehåller mycket information om data är det viktigt att kontrollera vilka delar av datakatalogen som exponeras och till vem. Exempelvis kan beskrivningar av ingående datamängder vara något som delas medan information kring källor som används vid dataintegrering kan vara känslig.

Existerande lösningar för behörighetshantering (2/2)

DIGG

Myndigheten för Digital Förvaltning (DIGG) har som regeringsuppdrag 2025 att utveckla en sammanhållen infrastruktur för identitets- och behörighetshantering, som ska kunna tillhandahållas inom ramen för Ena – Sveriges digitala infrastruktur.

DIGG har många pågående initiativ för att stödja en fullständig identitets- och behörighetslöning. Bland annat har DIGG i sitt arbete utvecklat tjänsten Säker Digital Kommunikation (SDK) som möjliggör säker överföring av känslig information, såsom hälsodata, mellan offentliga aktörer. Ett annat exempel är auktorisationssystem för elektronisk identitet och digital post som lanserades i maj 2025¹.

Analys

I E-hälsomyndighetens förslag för Nationell Digital Infrastruktur (NDI) beskrivs att det fortfarande finns vägval att göra kring regeringsuppdraget att utveckla en sammanhållen infrastruktur för identitets- och behörighetshantering² vilket indikerar att lösningen inte i sin nuvarande form är redo att användas för den regiongemensamma datahubben.

Inera

Anropsbehörighet för tjänstekonsumenter regleras på flera nivåer och baseras på organisationstillit, där tillit etableras genom lagar, avtal och regelverk.

Organisationstillit innebär förtroende för en aktörs förmåga att hantera information enligt gällande lagar, informationssäkerhet (t.ex. ISO27001) och kvalitetskrav, samt att de har ett offentligt finansierat uppdrag och relevanta försäkringar. Tillitskedjor säkerställer att varje aktör i ett digitalt informationsutbyte upprätthåller tillit genom att skydda kommunikation (t.ex. med TLS), validera identiteter och följa åtkomstpolicy.

Åtkomsthantering kan realiseras med tekniker som OAuth2 eller klientcertifikat, medan informationsbehörighet säkerställer att endast behörig information delas, enligt lokala regelverk och policy.

Slutsats

- Vid en implementation bör en kartläggning över tillgängliga och efterfrågade behörighetslösningar utföras, både mellan användare och system samt system sinsemellan.
- Samverkan bör etableras och en dialog bör föras med DIGG och E-hälsomyndigheten för att djupare kunna analysera om, hur och när deras lösning för identitets- och behörighetshantering kan nyttjas för en regiongemensam datahubb.

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

1. [Digitala tjänster | Digg](#)

2. [genomforandet-av-en-nationell-digital-infrastruktur---slutredovisning.pdf](#)

Metatahantering och datakatalog

Vad är metadata och metadatahantering?

Om metadata

Metadata är data som beskriver och ger information om annan data. Det fungerar som en typ av "data om data" och hjälper till att organisera, identifiera, och hantera information. Det finns flera olika typer av metadata. Några av dessa beskrivs nedan.

- **Beskrivande metadata:** Används för att beskriva innehållet i en resurs, inklusive titel, författare, ämnesord och beskrivningar.
- **Strukturell metadata:** Beskriver datas organisation och relation mellan olika delar.
- **Administrativ metadata:** Hantera datatillgångar, inklusive åtkomst, rättigheter och tekniska detaljer.
- **Kvalitetsmetadata:** Beskriva datas kvalitet och tillförlitlighet.
- **Semantisk metadata:** Ger betydelse och kontext till data genom att definiera relationer mellan datakomponenter och använda standardiserade termer.
- **Teknisk metadata:** Beskriver tekniska aspekter av en resurs, filformat, datastorlek, skapandedatum, uppdateringsfrekvens, lagringsplats och tekniska krav (t.ex. kompatibla verktyg).
- **Operationell metadata:** Beskriver processer relaterade till datahantering, inklusive insamling, bearbetning och användning, samt dokumenterar prestanda, ändringar och säkerhetsåtgärder.

Sammanfattningsvis finns det många olika sätt att beskriva och gruppera metadata men oavsett typ av metadata är metadata **avgörande för att effektivt hantera och navigera i stora mängder data**, vilket gör det enklare för användare och system att förstå och använda informationen.

Utvalda koncept

1. **Data Lineage:** Spårar datas ursprung, förflyttningar och transformationer när data flyttar genom och mellan system.
2. **Data Discovery:** Gör data spår- och sökbar och möjliggör att användare kan hitta relevant information och dataset baserat på metadata, taggar eller användning.
3. **Data Governance & Data Management:** Genom arbete med data governance och data management kan nödvändiga processer och roller som behövs för att berika information i en datakatalog etableras. Exempelvis genom att spåra dataäggande, data stewards, behörighetspolicies eller säkerhetsklasser.

Om metadatahantering

Metadatahantering innebär processen att skapa, organisera, lagra, uppdatera och använda metadata. Syftet är att göra information lättare att hitta, förstå och använda. Här är en kort översikt:

- **Skapa metadata:** Definiera och registrera information om data som titel, författare, datum, format eller nyckelord.
- **Organisera:** Strukturera metadata med standarder (t.ex. Health DCAT-AP) för att säkerställa konsekvens och sökbarhet.
- **Lagra:** Spara metadata i databaser, kataloger eller inbäddat i filer.
- **Uppdatera:** Hålla metadata aktuell vid förändringar i data.
- **Använda:** Underlätta sökning, analys, delning och interoperabilitet mellan system.

Vad är datakatalog och metadatastandard?

Om datakatalog

En datakatalog är verktyg som **organiserar en samling av metadata** som beskriver och ger information om olika datakällor, dataset eller resurser inom en viss domän eller organisation.

Användare

Genom att använda en datakatalog kan användare enkelt söka efter och få tillgång till specifika dataset eller resurser. Katalogen fungerar som en guide som gör det lättare att navigera i stora mängder data.

System

Datakataloger underlättar datadelning och interoperabilitet mellan olika system och plattformar genom att standardisera hur data beskrivs och dokumenteras. Detta är särskilt viktigt i sammanhang där flera aktörer samarbetar och behöver förstå varandras data.

Utvalda koncept

1. **Dataupplysningar:** Information om datakällor, inklusive ursprung, format och användning.
2. **Dataklassificering:** Kategorisering av data baserat på dess typ, känslighet och användning.
3. **Användargränssnitt:** Verktyg och funktioner som gör det möjligt för användare att enkelt navigera och söka i datakatalogen.

Om metadatastandard

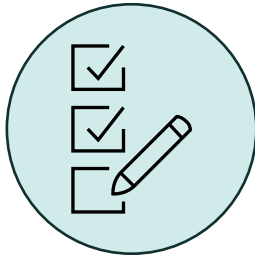
En metadatastandard eller metadata-specifikation är en uppsättning regler och riktlinjer som definierar hur metadata ska struktureras, beskrivas och hanteras i en datakatalog.

- Metadata-specifikationer ger en tydlig vägledning för hur metadata ska dokumenteras och hanteras, vilket gör det enklare för användare att förstå och navigera i datakataloger.
- Genom att använda en gemensam metadata-specifikation kan olika datakataloger kommunicera och dela information mer effektivt.

Metadatahantering, metadatastandard och datakatalog

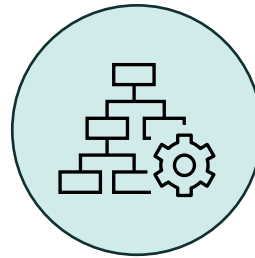
Nedan presenteras hur allt hänger ihop

Metadatastandard



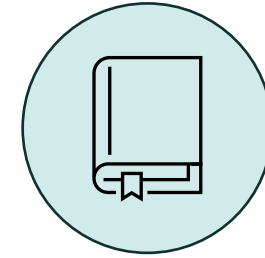
Metadatastandarder beskriver hur metadata ska struktureras och formateras.

Metadatahantering



Metadatahantering använder metadatastandarder för att samla in, organisera och underhålla metadata på ett strukturerat sätt. Utan standarder riskerar metadata att bli inkonsekvent eller svåransvänd.




Datakatalog



Datakatalogen bygger på metadatahantering och drar nytta av metadatastandarder för att presentera metadata på ett användarvänligt sätt. Katalogen gör det möjligt för användare att söka och utforska data baserat på standardiserad metadata.

EHDS, myndighetsfördelning och regioner

Nedan beskrivs hur myndigheter samt regioner kommer behöva förhålla sig till metadatastandard, metadatahantering och datakatalog baserat på EHDS

EHDS	 Metadatastandard	 Metadatahantering	 Datakatalog
	Health DCAT-AP är en utvidgning av DCAT-AP, specifikt anpassad för hälsodata inom ramen för EHDS. Den är utformad för att standardisera metadata för att förbättra interoperabilitet, sökbarhet och delning av hälsodata över EU:s datarum.	Medlemsstater måste implementera system för att samla in, lagra, underhålla och uppdatera metadata enligt Health DCAT-AP. Metadatahanteringssystem ska stödja bland annat datakvalitet, spårbarhet och dataskydd.	Medlemsstater måste upprätta nationella datakataloger som listar hälsodamängder och deras metadata enligt Health DCAT-AP. Katalogerna ska integreras med den EU-gemensamma plattformen HealthData@EU för att möjliggöra gränsöverskridande datadelning. Datakataloger måste vara sökbara, användarvänliga och stödja maskinläsbara metadata.
	Myndighetsfördelning¹ Health DCAT-AP är inte implementerat i större utsträckning i Sverige. Utredningen föreslår att Socialstyrelsen, i samarbete med Vetenskapsrådet, ska utveckla en nationell datasetkatalog baserat på Health DCAT-AP, med utgångspunkt i Vetenskapsrådets datakatalog RUT (Register Utiliser Tool)	Utredningen föreslår att Socialstyrelsen som samordnande HDAB ska hantera dataanmälningar, datatillstånd, datainsamling och anonymisering/pseudonymisering, vilket inkluderar metadatahantering. SCB och Vetenskapsrådet ska utreda säkra behandlingsmiljöer.	Utredningen föreslår att Socialstyrelsen, i samarbete med Vetenskapsrådet, ska tillhandahålla en nationell datakatalog baserat på RUT.
	Regioner Regionerna kommer troligtvis behöva implementera Health DCAT-AP eller kompatibel standard med mappning och anpassa existerande system och datalager för att generera metadata enligt Health DCAT-AP.	Regionerna kommer behöva stärka den existerande metadatahanteringen för att möta kraven från EHDS.	Regionerna kommer behöva bidra med Health DCAT-AP-metadata till den nationella katalogen antingen genom en egen datakatalog eller genom automatiserade flöden.

Nationellt system för kunskapsstyrning
Hälso- och sjukvård

1. [det-europeiska-halsodataområdet--sekundaranvandning-och-ansvarsfordelning-mellan-myndigheter.pdf](#)

Slutsats

Nedan beskrivs hur myndigheter samt regioner kommer behöva förhålla sig till metadatastandard, metadatahantering och datakatalog baserat på EHDS



Metadatastandard



Metadatahantering



Datakatalog

Slutsats

- Med existerande kravställning kring Health DCAT-AP är projektets slutsats att regionerna bör börja anpassa existerande metadatainsamling i enlighet med Health DCAT-AP, i den utsträckning det är möjligt.

Slutsats

- Gällande metadatahantering finns många fördelar med att ha en effektiv metadatahantering. I delprojekt 2, Analys av regionernas dataförmågor, föreslås ett strukturerat arbete kring Data Governance och Data Management där metadatahantering är en komponent.
- Projektet föreslår samarbete med Vetenskapsrådet för att gemensamt detaljera regionalt arbete med metadatahantering.

Slutsats

- Projektet föreslår att regionerna etablerar regionala datakataloger som kan federeras till en regiongemensam katalog som i sin tur är kompatibel med den nationella datakatalogen.
- Projektet föreslår samarbete med Vetenskapsrådet för att säkerställa kompatibilitet mellan regionernas datakataloger, den regiongemensamma datakatalogen och den nationella datasetkatalogen

Slutsats

- Projektet föreslår att metadatastandard, metadatahantering och datakatalog inkluderas i en framtida demonstration av referensarkitektur

Existerande datainsamlingar

Existerande datainsamlingar

Hälsodataregister

- Hälsodataregister är systematiska databaser som samlar in information om vård, tandvård och läkemedel som befolkningen fått i Sverige. Socialstyrelsen ansvarar för de nationella hälsodataregistren.
- De används för att analysera och förbättra hälso- och sjukvården samt för forskning och innovation och innehåller uppgifter om diagnoser, behandlingar, läkemedelsanvändning och patientresultat, vilket möjliggör en djupgående analys av vårdkvalitet och patientnytta.
- Uppgifter skyddas av lag och kan endast användas för forskning efter noggrann prövning, ofta i anonymiserad form.

Kvalitetsregister

- Kvalitetsregister i Sverige är individbaserade databaser som samlar in information om diagnoser, behandlingar och resultat för patienter.
- Det finns över 150 nationella kvalitetsregister som används för att utveckla och förbättra hälso- och sjukvården, samt som värdefulla källor för forskning och innovation.
- Kvalitetsregistren ger verksamheter återkoppling på sina resultat och möjliggör jämförelser mellan olika regioner och vårdgivare, vilket stödjer analys- och förbättringsarbete.
- Kvalitetsregistren är en del av det nationella systemet för kunskapsstyrning inom hälso- och sjukvård.

Andra register

SKR samordnar flera datainsamlingar. Några av dessa är:

Kostnad per patient (KPP): Tillhandahåller kostnader för vårdkontakter uppdelat per region, diagnos och sjukhus, samt övrig vård- och patientinformation.

Ekonomi- och verksamhetsstatistik: Innehåller uppgifter om verksamhet och ekonomi i regioner för att ge en nationell översikt och underlätta jämförelser över tid.

Nationell patientenkät: Samlar in patienters upplevelser av vården. Resultaten används för ledning och styrning samt möjliggör jämförelser.

Vilket arbete behövs för att underlätta skick till nationella register?

Kartläggning av variabler: Det är viktigt att identifiera och sammanställa alla variabler som ska rapporteras till de olika registren. Detta inkluderar att förstå specifikationerna för varje variabel och hur de ska rapporteras.

Standardisering av begrepp: En gemensam informatik- och datamodell behövs för att säkerställa att alla regioner använder samma definitioner och begrepp. Detta minskar risken för olika tolkningar och förbättrar jämförbarheten.

Teknisk lösning: Utveckling av tekniska lösningar som kan automatisera överföringen av data till registren.

Kontinuerlig uppdatering och anpassning: Eftersom specifikationer och krav kan förändras över tid, måste system och processer kontinuerligt uppdateras och anpassas för att möta nya krav.

Slutsats

- Datahubben kan över tid bli kanalen från regionerna till de existerande datainsamlingarna.
- Regionerna skickar dataunderlag till en plats, datahubben, som i sin tur behandlar data och skickar till datainsamlingarna
- Se även avsnitt 5 kring realisering av användningsfall

**Nationellt system
för kunskapsstyrning
Hälso- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Cloud vs. hybrid vs. on-prem

Jämförelse mellan cloud, hybrid och on-prem (1/2)

Det finns olika alternativ för hur en organisation kan välja att installera, administrera och underhålla sina programvaror och IT-system. För datahubben är de olika alternativen samt de framtida vägvalen centrala utifrån hur delar av referensarkitekturen kommer att realiseras.

	Cloud	Hybrid	On-prem
Beskrivning	Cloud innebär att hårdvaran som används för att genomföra olika operationer ägs och underhålls av en extern part.	Hybrid innebär en kombination av cloud och on-premise-lösningar, där man kan hyra serverutrymme och installera egen programvara.	On-premises (on-prem) innebär att den hårdvara som en organisation använder för sina IT-lösningar är placerad inom organisationens egna lokaler, och att organisationen ansvarar för underhållet.
Skalbarhet	Det är enkelt att skala upp och ner resurserna baserat på efterfrågan, eftersom ingen ytterligare hårdvara behövs.	Flexibel skalbarhet; enkelt att justera resurser hos leverantören, men kan kräva investeringar i egen hårdvara.	Uppskalning kan kräva att ytterligare hårdvara köps in och installeras, vilket kan ta tid.
Kostnad	Lägre uppstartskostnad eftersom cloud ofta använder en "betala för vad du använder"-modell och att den nödvändiga hårdvaran redan existerar.	Kostnaderna varierar; initialt kan det vara lägre än on-premise, men inkluderar hyra och investeringar.	Det krävs stora investeringar för hårdvara, mjukvara och infrastruktur.
Tillgänglighet	Data och applikationer kan användas överallt där det finns en internetuppkoppling.	Data och applikationer är tillgängliga både lokalt och via internet, vilket kräver korrekt konfiguration.	Tillgång till on-prem-system utanför det befintliga nätverket kräver ytterligare konfiguration och säkerhetsåtgärder.
Säkerhet	Lagring av känslig data utanför företagets egna servrar kan skapa säkerhets- och complianceproblem beroende på existerande krav.	Säkerheten kräver noggrant övervägande; känslig data kan lagras både lokalt och i molnet.	Det finns en förstärkt säkerhet för känslig data, vilket minimerar riskerna kopplade till tredjepartstillgång.
Kontroll	Användare har mindre kontroll över infrastrukturen, och olika leverantörer erbjuder olika tjänster, vilket kan medföra en inlåsningsseffekt.	Blandad kontroll; användare kan styra egen programvara men är beroende av leverantörens infrastruktur.	Organisationen har full kontroll över hårdvara, mjukvara och data.
Prestanda	Prestanda beror på internettillgång och leverantörens infrastruktur.	Prestandan varierar; vissa applikationer kan prestera bättre lokalt, medan andra drar nytta av cloud-resurser.	Vissa typer av applikationer kan prestera bättre eftersom data hanteras lokalt.
Compliance	Compliance är mer komplext eftersom det finns stora beroenden av hur leverantören hanterar data.	Compliance är mer hanterbart än i ren cloud-lösning, men det finns krav att följa leverantörens riktlinjer.	Compliance-krav är lättare att uppfylla eftersom organisationen har full kontroll över alla komponenter.

**Nationellt system
för kunskapsstyrning
Hälso- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Jämförelse mellan cloud, hybrid och on-prem (2/2)

Marknadstrender kring cloud

Sedan tillträdet av den nya amerikanska administrationen i januari 2025, har flera europeiska myndigheter, regeringar och organisationer valt att aktivt leta efter alternativ till amerikanska cloud-leverantörer. Eftersom en stor del av marknaden domineras av de amerikanska leverantörerna Amazon Web Services (AWS), Microsoft Azure och Google Cloud försöker nu flera aktörer leta efter andra lösningar för minska beroendet av dessa.

I Nederländerna har parlamentet godkänt en motion som uppmanar regeringen att minska sitt beroende till amerikanska mjukvarubolag. Lagstiftarna hänvisade främst till farhågor kring autonomi och cybersäkerhet.

Försäkringskassan har sedan 2017 i uppdrag av regeringen att erbjuda andra myndigheter samordnad och säker it-drift. Försäkringskassan använder en standardiserad modell för att ansluta myndigheter som inkluderar fyra faser: Initiera, planera, paketera samt förvalta och utveckla. Vid författandet av denna rapport är ett antal myndigheter anslutna till Försäkringskassans erbjudande, däribland E-hälsomyndigheten och Pensionsmyndigheten.

Nuläge kring hybrid och on-prem

Det finns få datacenter som helt drivs i egen regi av regioner eller andra närliggande aktörer (ex. Inera). Det vanligare alternativet är att leverantörer som Orange och TietoEvery ansvarar för datacentren medan regionerna väljer och ansvarar för programvara.

Diskussion

De rådande osäkerheterna kring framför allt cloud-frågeställningen, tillsammans med komplexiteten och storleken för den regiongemensamma datahubben, gör att renodlade cloud-lösningar inte är det primära alternativet baserat på nuvarande tillgänglig information.

Att använda befintliga helägda regionala datacenter begränsar frågan kring vem som skulle kunna vara värd för den regiongemensamma datahubben men det är ett alternativ beroende på utfallet i frågan kring vem som ska ansvara för datahubben.

Slutsats

- I nuläget lämpar sig en hybrid eller on-prem modell beroende på vem som i framtiden ska förvalta och husera de tekniska komponenterna i datahubben.
- Innan det beslutas om vem som i framtiden ska förvalta och husera tekniken bör en analys av kandidaternas existerande infrastruktur genomföras.

Säkra behandlingsmiljöer

Säkra behandlingsmiljöer (1/2)

Om säkra behandlingsmiljöer

En säker behandlingsmiljö beskriver hur en kontrollerad processeringsmiljö med hjälp av processer (program eller aktiviteter) kan utföras med utökad säkerhet för att skydda känslig data, säkerställa korrekt behörighet och förbygga felaktig användning av information. Detta kan jämföras med en låst och övervakat rum där en dator utför en uppgift. Bara specifika personer (eller program) har tillgång, allt i rummet är privat och skyddat och det finns specifika regler som beskriver vad som är tillåtet att göra i rummet.

Utvalda koncept

- **Behörighetshantering:** Säkra behandlingsmiljöer har en strikt behörighetshantering så enbart användare med tillstånd kan interagera med miljön.
- **Isolering:** Miljön och processerna inom miljön är isolerade från andra miljöer och verktyg. Bara godkända aktiviteter kan genomföras inom miljön.
- **Dataskydd:** Data i vila och rörelse är krypterat och i många fall är även data som kommer in i de säkra behandlingsmiljöerna pseudonymiserad eller anonymiserad.
- **Monitorering:** Samtliga aktiviteter och processer inom en säker behandlingsmiljö övervakas och loggas för att säkerställa att bara godkända användare och aktiviteter genomförs.

TEHDAS2 och EHDS

TEHDAS2 (Joint Action Towards European Health Data Space 2) är ett EU-projekt som ska skapa vägledningar och tekniska specifikationer för att underlätta införandet av det europeiska hälsodataområdet. Projektet omfattar 29 EU-länder och bygger vidare på tidigare arbete från TEHDAS1 och andra europeiska projekt.

TEHDAS2 har i februari 2025 avslutat en första remissomgång där *"Vägledning för dataanvändare om användning av data i en säker bearbetningsmiljö"* har publicerats. I en kommande remissomgång som går till konsultation september-oktober 2025 kommer bland annat följande remisser presenteras¹:

- *Guideline for Health Data Access Bodies on the procedures and formats for data access*
- *Technical specification for Health Data Access Bodies on the implementation of the common IT infrastructure*
- *Guideline for Health Data Access Bodies on linkage of health datasets*

Utfallet från dessa remisser kommer ha stor påverkan på hur den regiongemensamma datahubben ska hantera säkra behandlingsmiljöer.

Säkra behandlingsmiljöer (2/2)

SCB – MONA

MONA (Microdata Online Access) är Statistiska Centralbyråns (SCB:s) plattform för tillgängliggörande av mikrodata. I MONA kan användare göra bearbetningar via internet utan att mikrodata lämnar SCB. Data lagras i SCB:s databaser, behandlas, pseudonymiseras och tillgängliggörs i windowsmiljö via fjärruppkoppling¹.

Mikrodata från SCB:s register är främst för forsknings- och statistikändamål och i dagsläget finns register inom följande ämnesområden:

- Arbetsmarknad
- Befolkning
- Fastighetsregister
- Företagsregister och företagsundersökningar
- Hushållens ekonomi och utgifter
- Inkomst och taxering
- Levnadsförhållanden
- Demokrati
- Utbildning

Det finns inga specifika register eller tillhörande säkra behandlingsmiljöer för hälsodata.

Pågående lagstiftning i Sverige

Den 9:e oktober 2024 publicerades en utredning med uppdrag att möjliggöra en nationell infrastruktur. I utredningen analyseras och föreslås i promemorian de myndigheter som ska ansvara för tillgången till hälsodata samt vara nationell kontaktpunkt för sekundäranvändning av hälsodata.

I utredningen föreslås att Statistiska centralbyrån (SCB) ska vara den HDAB som ska ha ansvar för att tillhandahålla en eller flera säkra behandlingsmiljöer. Utöver det föreslås att SCB bör få i uppdrag att utreda hur systemet med säkra behandlingsmiljöer bör sättas upp och utvecklas. Uppdraget bör genomföras tillsammans med Vetenskapsrådet².

Slutsats

- Då det pågår europeiska utredningar genom TEHDAS2 kring säkra behandlingsmiljöer som kommer påverka hur dessa ska användas är projektets slutsats att nästa konsultationsrunda från TEHDAS2 i september bör inväntas.
- Gällande utvecklingen av säkra behandlingsmiljöer i Sverige bör ett samarbete med Statistikmyndigheten, SCB, upprättas för att säkerställa att framtida kravställning kring säkra behandlingsmiljöer kan mötas av datahubben.
- Fortsatt utredning om vilka andra säkra behandlingsmiljöer som existerar i Sverige.

1. [MONA – SCB:s plattform för mikrodata](#)

2. Det europeiska hälsodataområdet – sekundäranvändning och ansvarsfördelning mellan myndigheter, [det-europeiska-halsodataområdet--sekundaranvandning-och-ansvarsfordelning-mellan-myndigheter.pdf](#)

Containrar

Vad är containrar och hur används de?

Om containrar

Containrar är portabla enheter som används för att paketera en applikation inklusive allt som applikation behöver för att fungera. Detta inkluderar kod och bibliotek.

Ett sätt att se på containrar är att se det som en mindre version av en virtuell maskin där den stora skillnaden är att virtuella maskiner har ett operativsystem medan containrar använder sin värds operativsystem.

Vanliga användningsfall för containrar

- **Mikrotjänstarkitektur** – genom att dela upp tjänster och funktioner mellan olika containrar kan olika delar av arkitekturen skalas och driftsättas oberoende av varandra.
- **Data pipelines** – generellt innehåller dataflöden många verktyg med komplexa relationer som behöver konfigureras. Containrar kan se till att logik är portabel och paketerad vilket gör att komplexa dataflöden kan delas upp.
- **AI/ML** – AI och maskininlärning kräver ofta olika versioner av drivrutiner, bibliotek och miljöer. Genom att dela upp det i containrar kan varje modell köras oberoende av andra.
- **Olika miljöer** – Genom containrar kan samma applikation användas oberoende om miljön är cloud, on-prem eller en hybrid

Varför ska containrar appliceras i en regiongemensam miljö?

Tre huvudsakliga anledningar har identifierats

Säkerställa att data modelleras enhetligt

Eftersom enhetlighet är en förutsättning för att fullt kunna realisera värde från den tänkta arkitekturen betyder det att regionerna kommer behöva modellera sin data till samma gemensamma modell. Genom att använda ett modelleringsverktyg genom en container där slutmodellen är bestämd ökar förutsättningar för att data modelleras enhetligt.

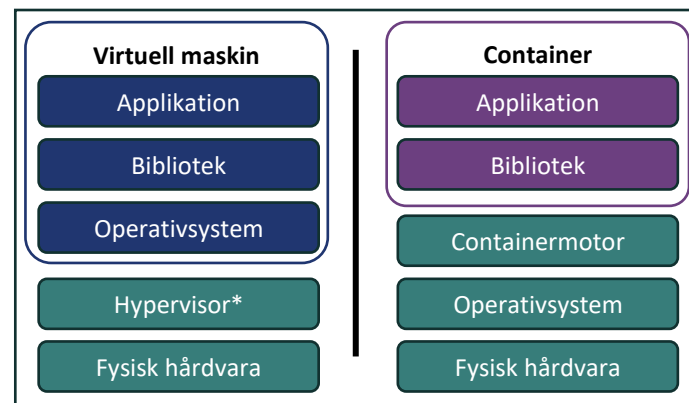
Minimera arbetsbördan på regionerna

Genom skapa container konfigurerings s.k. "Images" centralt så kan mycket tid sparas på regional nivå genom att samma centrala image används som ritning för att skapa containrar hos regionerna

Låta regionerna behålla full kontroll över sin data

Eftersom containrar installeras på regionala miljöer kommer regionerna ha full kontroll över vilka transformationer och modelleringar som görs på deras data.

Illustration över skillnader mellan virtuell maskin och containrar



**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

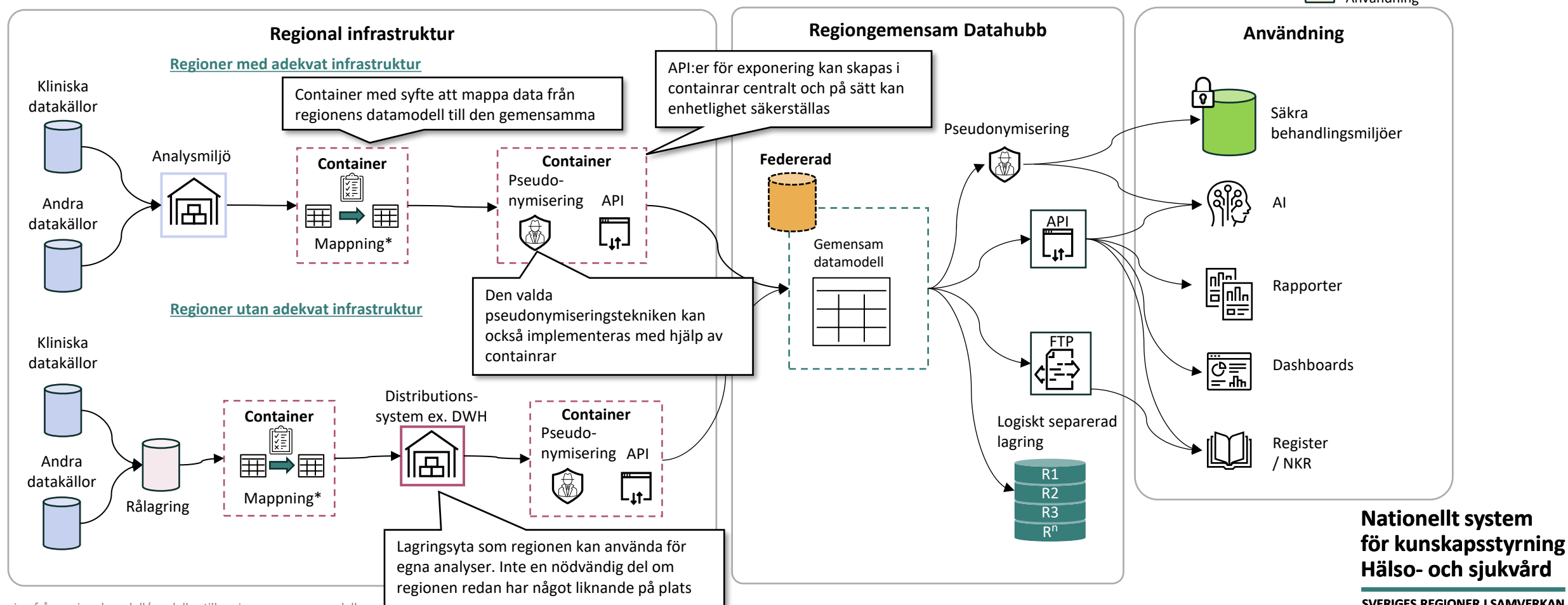
SVERIGES REGIONER I SAMVERKAN

* Hypervisor, även kallad Virtual Machine Monitorering (VMM), är en mjuk- eller hårdvara som möjliggör skapande och hantering av virtuella maskiner.

Exempel på containrar i realiseringsarkitekturen

I realiseringsarkitekturen kan containrar användas för att bygga upp en modulbaserad miljö. Containrar kan ses som byggklossar med specifik funktionalitet som sen kan appliceras och implementeras i regioner baserat på den regionens specifika behov. Nedan syns ett exempel hur det skulle kunna se ut.

- Befintlig regional funktionalitet
- Utökad regional funktionalitet
- Regiongemensam funktionalitet
- Användning





4.4 Summering slutsatser

Summering slutsats

Följande avsnitt summerar slutsatser som beskrivs under fördjupningsområdena. Den tredje kolumnen "Rekommendation" refererar till rekommendationerna som beskrivs under kapitel 7 och hur dessa hänger ihop med slutsatserna nedan. Alla fördjupningsområden har inte slutsatser som leder till rekommendationer och nästa steg.

Fördjupningsområde	#	Slutsats
Central lagring eller federering	S1	<ul style="list-style-type: none">Federering lämpar sig väl som lösning för datahubben. Genom federering kan regionernas existerande infrastrukturer, många med distributionssystem, nyttjas. Central lagring kan sedan användas efter den federerade för specifika datamängder där en batchhämtning är mer lämplig.
Datamodellering	S2	<ul style="list-style-type: none">För att möjliggöra fler analyser och samverkan mellan regioner behöver data vara mappad till en gemensam eller hybrid datamodell där en hybrid modell enbart skulle appliceras om en region har unika datamängder.Arbetet att mappa från regionernas egna modeller till den gemensamma modellen bör ske hos regionerna med stöd från kompetenscentrum för att regionerna ska behålla kontroll över data samt att modelleringsarbetet ska kunna underhållas över tid.
Datasäkerhet	S3	<ul style="list-style-type: none">En datakonsekvens- och riskanalys bör genomföras i ett tidigt skede.I uppstartsfasen under en implementation bör ett arbete ske kring datasäkerhet för att undersöka exempelvis nätverkssäkerhet.
Pseudonymisering	S4	<ul style="list-style-type: none">Olika pseudonymiseringstekniker kan appliceras för att uppnå tänkt syfte i den regiongemensamma datahubben.En juridisk utredning behöver genomföras för att tydliggöra vilka krav finns på pseudonymisering och anonymisering av data. Baserat på resultatet av utredningen kan ett beslut fattas kring vilken teknik som lämpar sig bäst i vårddatahubben.
Behörighetshantering	S5	<ul style="list-style-type: none">Vid en implementation bör en kartläggning över tillgängliga och efterfrågade behörighetslösningar utföras, både mellan användare och system samt system sinsemellan.Samverkan bör etableras och en dialog bör föras med DIGG och E-hälsomyndigheten för att djupare kunna analysera om, hur och när deras lösning för identitets- och behörighetshantering kan nyttjas för en regiongemensam datahubb.

Summering slutsats

Fördjupningsområde	#	Slutsats
Metadatahantering och datakatalog	S6	<ul style="list-style-type: none"> Med existerande kravställning kring Health DCAT-AP är projektets slutsats att regionerna bör börja anpassa existerande metadatainsamling i enlighet med Health DCAT-AP, i den utsträckning det är möjligt. Gällande metadatahantering så finns många fördelar med att ha en effektiv metadatahantering. I delprojekt 2 föreslås ett strukturerat arbete kring Data Governance och Data Management där metadatahantering är en komponent. Projektet föreslår samarbete med Vetenskapsrådet för att gemensamt detaljera regionalt arbete med metadatahantering. Projektet föreslår att regionerna etablerar regionala datakataloger som kan federeras till en regiongemensam katalog som i sin tur är kompatibel med den nationella datakatalogen. Projektet föreslår samarbete med Vetenskapsrådet för att säkerställa kompatibilitet mellan regionernas datakataloger, den regiongemensamma datakatalogen och den nationella katalogen. Projektet föreslår att metadatastandard, metadatahantering och datakatalog inkluderas i en framtida demonstration av referensarkitektur.
Kvalitetsregister och andra register	S7	<ul style="list-style-type: none"> Datahubben kan över tid bli kanalen från regionerna till de existerande datainsamlingarna. Regionerna skickar dataunderlag till <u>en</u> plats, datahubben, som i sin tur behandlar data och skickar till datainsamlingarna Se även avsnitt 5 kring realisering av användningsfall
Cloud, hybrid och on-prem	S8	<ul style="list-style-type: none"> I nuläget lämpar sig en hybrid eller on-prem modell beroende på vem som i framtiden ska förvalta och husera de tekniska komponenterna i datahubben. Innan det beslutas om vem som i framtiden ska förvalta och husera tekniken bör en analys av kandidaternas existerande infrastruktur genomföras.
Säkra behandlingsmiljöer	S9	<ul style="list-style-type: none"> Då det pågår europeiska utredningar genom TEHDAS2 kring säkra behandlingsmiljöer som kommer påverka hur dessa ska användas är projektets slutsats att nästa konsultationsrunda från TEHDAS2 i september bör inväntas. Gällande utvecklingen av säkra behandlingsmiljöer i Sverige bör ett samarbete med Statistikmyndigheten, SCB, upprättas för att säkerställa att framtida kravställning kring säkra behandlingsmiljöer kan mötas av datahubben. Fortsatt utredning om vilka andra säkra behandlingsmiljöer som existerar i Sverige.

5

Analys av användningsfall

Analys av användningsfall

I delprojekt 3, Användningsfall, identifierades ett antal prioriterade användningsfall för den framtida hälsodatalösningen. Dessa prioriterades utifrån låg komplexitet och högt värde, baserat på delprojektets erfarenheter och kompetens.

I detta avsnitt genomförs en analys av den framtagna referensarkitekturen utifrån de prioriterade användningsfallen. Syftet är att säkerställa att referensarkitekturen har utformats på ett sätt som effektivt möter användarnas behov och faktiskt ger mervärde i jämförelse med nuvarande lösningar.

Tillvägagångssätt

- Analysen genomfördes i form av workshops med en utvald referensgrupp av experter inom de olika användningsfallen. Den ägde rum under två tillfällen, där varje workshop varade i två timmar.
- Gruppen arbetade metodiskt genom ett användningsfall i taget och analyserade hela flödet i referensarkitekturen. Under diskussionerna fokuserade de på hur lösningen skulle fungera, vilka skillnader som fanns jämfört med nuvarande lösningar samt identifierade eventuella hinder och möjliga lösningar.
- Slutsatsen blev att samtliga tre användningsfall kan uppfyllas med referensarkitekturen, och att den regiongemensamma datahubben kommer att underlätta arbetet jämfört med dagens lösningar.

De prioriterade användningsfallen

1 Skapa en överblick av vårdplatser

2 Skapa en tillförlitlig nationell jämförelse av vårdköer och väntetider i samtliga regioner

3 Skapa underlag för kostnader på individnivå över tid

Användningsfall 1: Skapa en överblick av vårdplatser

I nedan exempel beskrivs hur användningsfall 1 skulle implementeras i den regiongemensamma datahubben

Skapa en överblick av vårdplatser

1. Skapa/applicera en definition av vad en **vårdplats** är inkl. ev. undergrupperingar
2. **Samla vårdplatsdata** från regionerna i regiongemensam vårddatahubb inklusive kontinuerlig uppdatering
3. Skapa en **visualisering** av tillgängliga vårdplatser

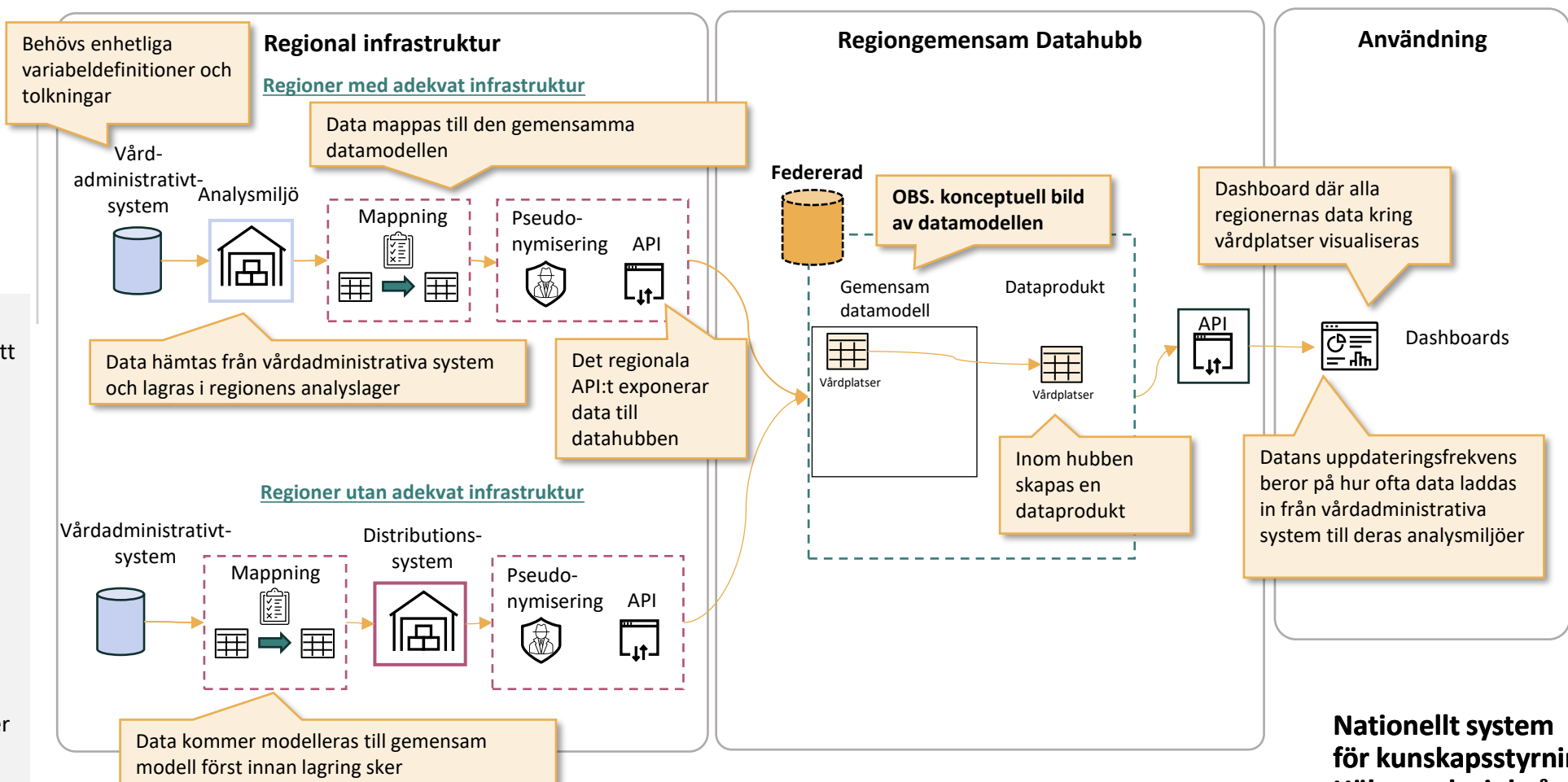
Analys av användningsfall 1

Förutsättningen för detta användningsfall är att ha enhetliga variabeldefinitioner.

Datahubben möjliggör att flera regioners överblick av vårdplatser kan visualiseras och jämföras med en hög frekvens.

Det förstnämnda kommer inte lösas av datahubben utan behöver ske inom ramarna för arbetet med variabelförbättring. Resterande värden kommer kunna realiseras genom datahubben.

Användarfall 1 kommer kunna realiseras under förutsättning att arbetet med enhetliga variabler och tolkningar genomförs.



**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

Användningsfall 2: Skapa en tillförlitlig nationell jämförelse av vårdköer och väntetider i samtliga regioner

I nedan exempel beskrivs hur användningsfall 2 skulle implementeras i den regiongemensamma datahubben

Skapa en tillförlitlig nationell jämförelse av vårdköer och väntetidsuppföljning i samtliga regioner

1. Skapa/applicera en **definition av vilket problem** en väntande patient väntar på vård för inkl. ev. undergrupperingar
2. Vidareutveckla **visualisering** av data.

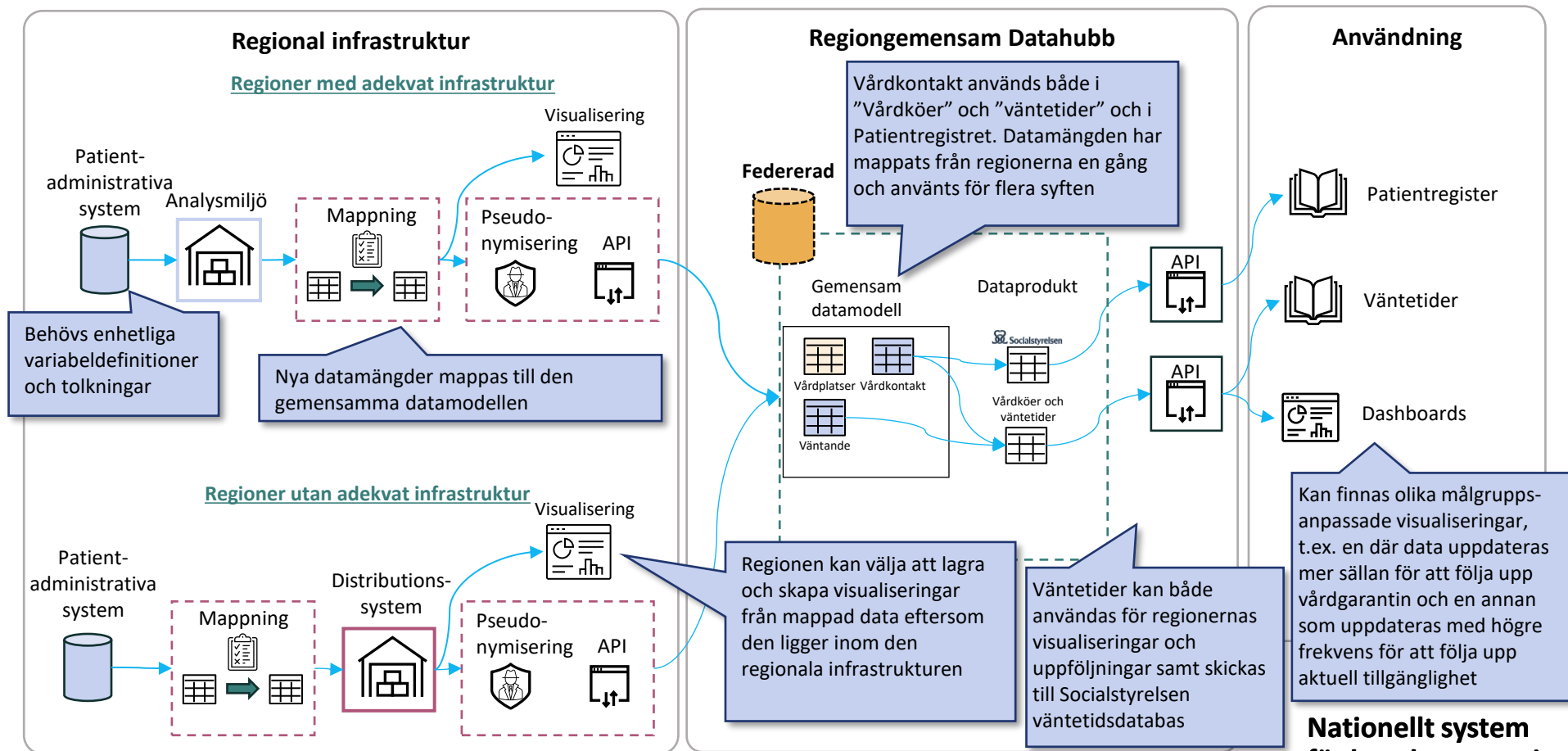
Analys av användningsfall 2

Förutsättningen för detta användningsfall är att ha enhetliga variabeldefinitioner.

Datahubben möjliggör Användningsfall 2. Gällande uppföljning av vårdgarantin kommer det finnas få utmaningar kring uppdatering och frekvens eftersom den typen av data kan uppdateras mer sällan (dagligen/per vecka). Frekvent uppföljning av aktuell tillgänglighet i vården ställer krav på hur ofta data laddas in i regionernas analysmiljöer.

När mer och olika data börjar finnas tillgänglig kan den stora nyttan av att ha en datahubb med en gemensam datamodell realiseras.

Datamängder som mappats och tillgänglig-gjorts en gång behöver inte göras om och kan nyttjas till flera syften.



Nationellt system för kunskapsstyrning Hälso- och sjukvård

SVERIGES REGIONER I SAMVERKAN

Användningsfall 3: Skapa underlag för kostnader på individnivå över tid

I nedan exempel beskrivs hur användningsfall 3 skulle implementeras i den regiongemensamma datahubben

Skapa underlag för kostnader på individnivå över tid

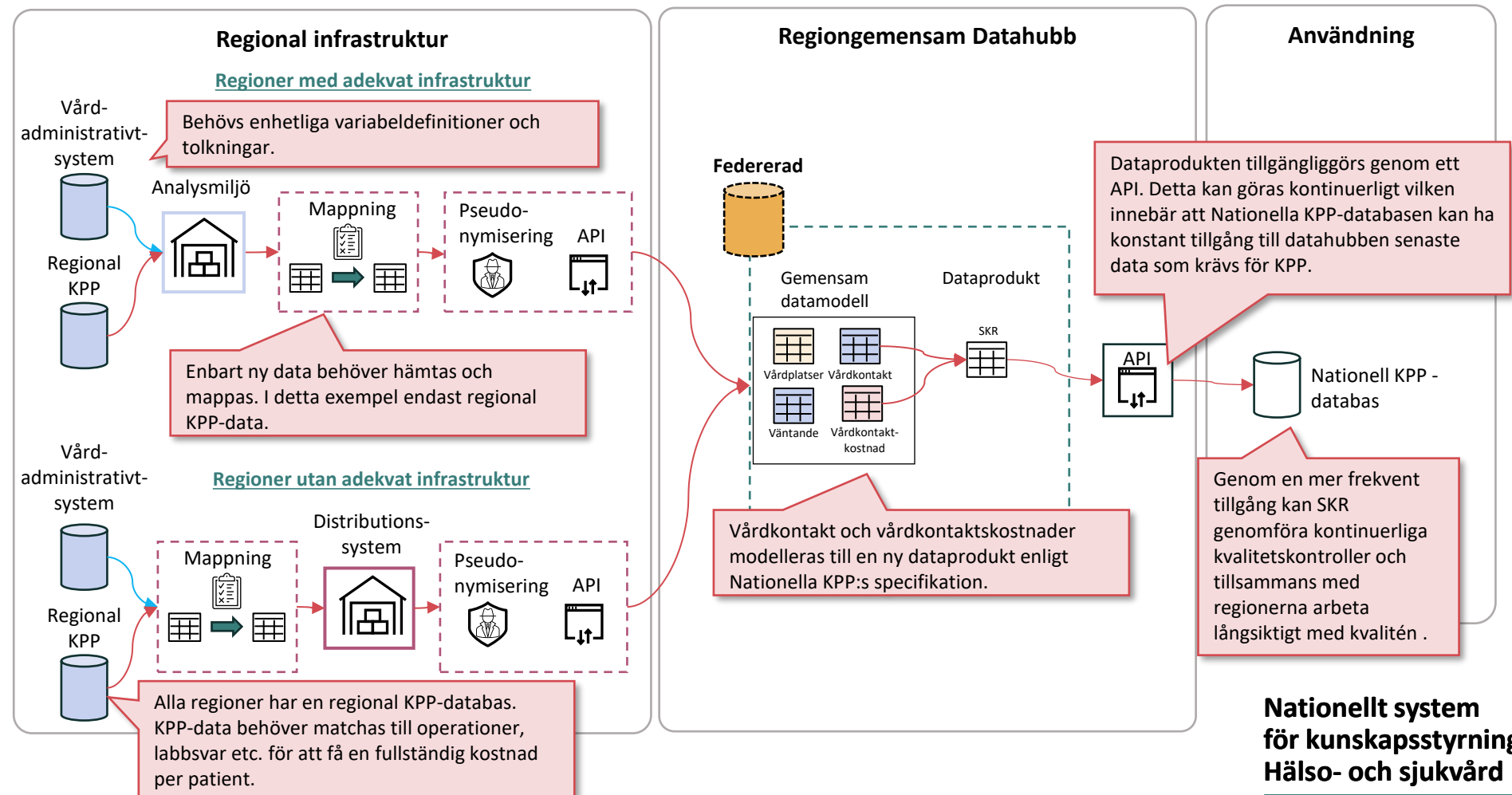
1. Skapa/applicera en **definition av kostnader** som kopplar till en enskild individ
2. **Samla in kostnadsdata** på individnivå och följ över tid
3. **Skapa underlag** där man kan följa kostnader på individnivå

Analys av användningsfall 3

Förutsättningen för detta användningsfall är att ha enhetliga variabeldefinitioner.

Eftersom viss data redan modellerats och tillgängliggjorts i datahubben i tidigare användningsfall så behöver bara arbetet utföras för regional KPP-data för detta användningsfall.

Likt tidigare användningsfall beror tillgänglighet på regionernas val av inhämtningsfrekvens i sina analysmiljöer. Användningsfallet kan realiseras i den nya arkitekturen och genom kontinuerlig tillgång samt samarbete kan ett mer långsiktigt kvalitetsarbete minska arbetsbördan med datakvalitet över tid.



**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN



6

Design av drift- och förvaltningsmodell

Drift- och förvaltningsmodeller

En teknisk lösning består av olika tjänster och komponenter som kan driftas och förvaltas på olika sätt.

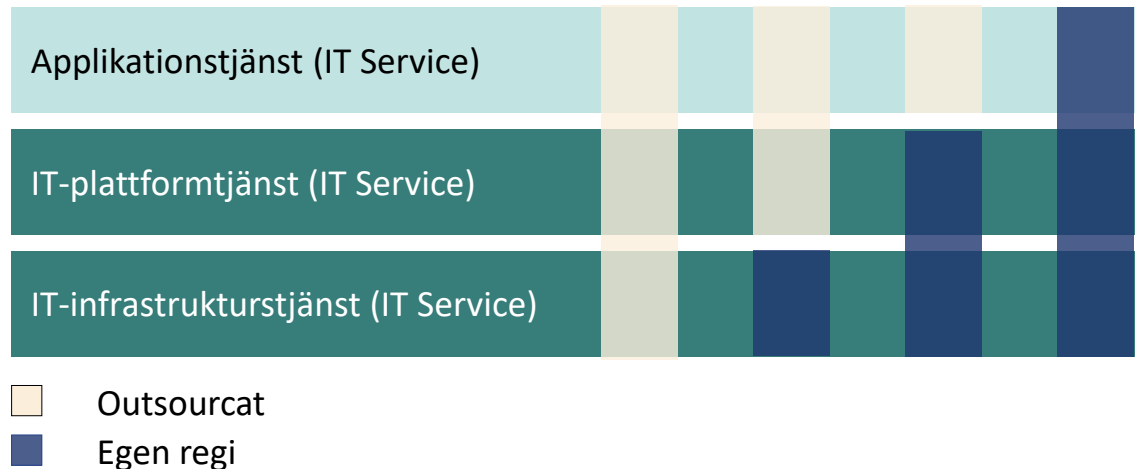
Nedan illustration beskriver en vanlig uppdelning av dessa komponenter.

	<i>Funktion</i>	<i>Exempel</i>
Applikationstjänst (IT Service)	<i>BI, analys, AI, beslutsstöd, ML-modeller</i>	<i>Power BI, Tableau, Qlik Sense</i>
IT-plattformstjänst (IT Service)	<i>Dataintegration, säkerhet, API:er, identitetshantering</i>	<i>Snowflake, Starburst, Keycloak</i>
 Data Management	<i>Dataflöden, datakvalitet, metadata, ETL</i>	<i>Talend, Informatica, Apache NiFi</i>
 Data Governance	<i>Policys, GDPR, åtkomstkontroll, compliance</i>	<i>Collibra, Alation, Apache Atlas</i>
IT-infrastrukturstjänst (IT Service)	<i>Lagring, nätverk, beräkning, moln</i>	<i>Serverar, datacenter, nätverk</i>

Drift- och förvaltningsmodeller

Nedan bild illustrerar hur den regiongemensamma vårddatahubben kan delas upp i olika IT-komponenter, där dessa kan hanteras i egen regi, outsourcat eller i kombination för respektive komponent.

Illustrativt exempel – inkluderade komponenter kan uppdateras efter att referensarkitekturen är färdigställd



Medskick vid val av modeller

Kombination av drift- och förvaltningsmodeller

- Eftersom det går att kombinera outsourcing och egen regi på många olika sätt är det just nu inte möjligt att lista och jämföra alla möjliga kombinationer
- Möjligheterna beror på hur vårddatahubben slutgiltigt utformas och designas samt ett antal strategiska inriktningar och principer

Design av drift- och förvaltningsmodeller

- Framgent är det lämpligt att arbeta hypotesbaserat för att utreda lämplig modell för respektive komponent av den framtida av arkitekturen där aspekter kopplat till paketering, behov samt prismässig påverkan behöver vägas samman
- Det finns ett värde i att designa en förvaltningsmodell där kombinationen av outsourcing och egen regi är flexibel och kan varieras över tid beroende på underliggande krav och behovsbild

Första steg och principer

- Innan utredning av drift- och förvaltningsmodell påbörjas är det lämpligt att definiera en övergripande strategi som blir ledande i kommande beslutsfattande för att:
 - Etablera referensarkitekturs **övergripande realiseringsstrategi**
 - Utveckla **lämplig governancemodell** baserat på referensarkitektur
 - Skapa förutsättningar för att **utreda kostnadsmässig och finansiell påverkan** på respektive föreslagen paketering

Upplägg för drift-och förvaltningsmodell kommer att se olika ut baserat på hur referensarkitekturen framgent paketeras och etableras

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Se kommande avsnitt 7, Rekommendationer och nästa steg

Finansiering

Finansieringsalternativ

Eventuellt för förstudie- och genomförande-projekt

Regionfinansiering	Statligt stöd	Avgiftsfinansiering	EU-finansiering genom EHDS och Horizon Europe	Vinnova, Vetenskapsrådet
<ul style="list-style-type: none"> Fördelas enligt storlek eller användning. Möjlighet att införa en avgiftsmodell där de som använder hubben mer betalar mer. Tilläggsfinansiering Övergångsperiod (omfördelning SKR) Alternativ för långsiktig och stabil finansiering. 	<ul style="list-style-type: none"> Staten kan bidra genom riktade bidrag för utveckling och säkerhet. Kan säkerställa att hubben håller nationella krav. Regeringen inrättar effektivitetsdelegation för hälso- och sjukvården - Regeringen.se 	<ul style="list-style-type: none"> Forskare och företag kan betala för tillgång till specifika dataset eller analysverktyg. kunskapsstyrning - industrisamverkan Alternativ för delfinansiering. 	<ul style="list-style-type: none"> Möjlighet att söka medel för specifika forsknings- och innovationsprojekt. Alternativ för kortsiktig finansiering. Möjliga vägar är EU4Health, Digital Europe, Cluster 1: Health och Horizon Europe. 	<ul style="list-style-type: none"> Möjlighet att söka medel för specifika forsknings- och innovationsprojekt. Alternativ för kortsiktig finansiering. Utllysningar av relevanta innovationsprojekt sker löpande. Exempel är Medtech4Health och internationellt samarbete för hälsoinnovation

Finansiering hos liknande initiativ*



Slutsats

Delprojektet har diskuterat ovan finansieringsalternativ för den regiongemensamma vårddatahubben. För vårddatahubben skulle olika alternativ kunna vara relevanta i och med att olika delar i hubben kan finansieras på olika sätt (förvaltning och drift, utveckling, utlämnande etc)

Nationellt system för kunskapsstyrning Hälso- och sjukvård

SVERIGES REGIONER I SAMVERKAN

* Mappningen av finansiering av liknande initiativ är inte att betrakta som fullständigt uttömmande givet att all information inte är offentlig

7

Rekommendationer och nästa steg

Delprojektets rekommendationer och nästa steg

Inledning till projektets rekommendationer

- Projektet har under genomförandet av delprojektet arbetat i enlighet med en struktur, där referensarkitekturen utgör den centrala och största komponenten och anses som delprojektets huvudsakliga leverabel.

Delprojektets **samlade rekommendation** är att under år 0 tillsätta nedan arbetspaket*

1 Arbetspaket 1: Teknisk demonstration och kravställning

Fortsatt **utreda, detaljera samt demonstrera** delar av den **föreslagna referensarkitekturen** samt ingående komponenter

2 Arbetspaket 2: Teknisk realisering

Utreda och besluta om hur **tekniska komponenter** i vårddatahubben ska etableras och hur det påverkar regionernas **organisation** och **arbetssätt**, samt behov av **investeringar**.

För att på ett strukturerat sätt kunna **följa upp** samt **paketera framförda rekommendationer** används nedan struktur

#	Område	Föreslagen rekommendation och nästa steg	Syfte med framförd rekommendation	Slutsatser
4.1	Detaljerings referensarkitektur	<ul style="list-style-type: none"> Baserat på framtagna referensarkitektur, detaljera och konkretisera utvalda delkomponenter med särskild teknisk, kommersiell eller organisatorisk komplexitet: Exempelområden för fördjupning: Mappning från regionala datamodeller, behörighetsstyrning, dataåtkomst, <u>pseudonymisering</u> samt utökad regiongemensam funktionalitet. 	Givet projektets komplexa omfattning kommer delkomponenter i den föreslagna referensarkitekturen behöva detaljeras. Detta för att på ett strukturerat sätt kunna ta höjd för kommande delsteg och möjliggöra ytterligare analys samt förankring.	S4, S5, S6

Beskrivning av område

Beskrivning av syftet med respektive rekommendation

ID på aktuellt område

Beskrivning av rekommendation och nästa steg

Hänvisar till slutsatser från kapitel 4

För att kunna **planera inför ett föreslaget genomförande** av framförda rekommendationer har även en **genomförandeplanering** utformats



Genomförandeplanering innefattar:

- Rekommendationer och nästa steg
- Tidplan samt estimat för antal resurser
- Innehåll för respektive arbetspaket

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Detaljer kring genomförande av respektive arbetspaket, inklusive förslag till rekommendationer och nästa steg, återfinns på nästkommande bilder

Delprojektets rekommendationer (1/4)

#	Område	Föreslagen rekommendation och nästa steg	Syfte med framförd rekommendation	Slutsatser*
4.1	Detaljeringsreferensarkitektur	<ul style="list-style-type: none"> Baserat på framtagna referensarkitektur, detaljera och konkretisera utvalda delkomponenter med särskild teknisk, investeringsmässig eller organisatorisk komplexitet: <ul style="list-style-type: none"> Exempelområden för fördjupning: Mappning från regionala datamodeller, behörighetsstyrning, dataåtkomst, pseudonymisering samt utökad regiongemensam funktionalitet. 	Givet projektets komplexa omfattning kommer delkomponenter i den föreslagna referensarkitekturen behöva detaljeras. Detta för att på ett strukturerat sätt kunna realisera kommande framförda rekommendationer samt möjliggöra ytterligare analys och förankring.	S4, S5, S6
4.2	Demonstration referensarkitektur	<ul style="list-style-type: none"> Validera föreslagen referensarkitektur tillsammans med en eller flera regioner i form av en Proof of Concept (POC). <ul style="list-style-type: none"> Genomföra en POC baserat på projektets realiserings exempel (se kapitel 4.2 i denna rapport). För att möjliggöra demonstration av referensarkitekturs funktion föreslås att en POC baseras på tillgängliga tekniska komponenter i så stor utsträckning som möjligt. 	<p>Givet projektets utformning samt dess organisatoriska spännvidd är en central aspekt för rekommendationen att referensarkitekturen samt arkitekturella komponenter kan demonstreras på ett konceptuellt plan.</p> <p>Aktiviteten fyller vidare en tydlig funktion i internt såväl som externt kommunikativt syfte för att påvisa samt validera tilltänkt funktionalitet.</p>	S1, S6
4.3	Vidareutveckla kravspecifikation	<ul style="list-style-type: none"> Utifrån befintliga designkriterier, ta fram en detaljerad och sammanhållen kravspecifikation. Säkerställa att arbetet följer en etablerad metodik, vilket underlättar integrering av tidigare resultat och ger en stabil grund för kommande faser. 	Denna rekommendation syftar till att skapa struktur inför kommande steg. En tydlig kravspecifikation är avgörande för en effektiv marknadsdialog och för att kunna bedöma projektets ekonomiska och tekniska omfattning.	

Realisering av rekommendationerna ovan, föreslaget inom ramen för genomförande av:

1 Arbetspaket 1: Teknisk demonstration och kravställning

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Hänvisar till slutsatser från kapitel 4 i denna rapport.

Delprojektets rekommendationer (2/4)

#	Område	Föreslagen rekommendation och nästa steg	Syfte med framförd rekommendation	Slutsatser*
4.4	Utveckling samt förankring av datamodell	<ul style="list-style-type: none"> • Utifrån tidigare arbete, definiera en gemensam datamodell. Detta säkerställer att data kan modelleras enligt vedertagen metodik i referensarkitekturen och möjliggör framtida analys och jämförelse. • Engagera/etablera regionala förankringsgrupperingar för att säkerställa strategisk inriktning, innehåll, förankring och utveckling av det nya arbetssättet. 	<p>En central aspekt för centraliserad datajämförelse är en gemensam datamodell. Detta område är komplext med tidigare regionala insatser som delvis skiljer sig åt. Rekommendationen syftar till att:</p> <ul style="list-style-type: none"> • Identifiera den minsta gemensamma nämnaren bland existerande regionala datamodeller. • Utvärdera och anpassa internationella standarder och modeller för en svensk kontext. • Utvärdera befintliga regionala datamodeller och välja den med bredast användning som utgångspunkt. 	S2
4.5	Fortsatt teknisk omvärldsbevakning och intressent-kartläggning	<ul style="list-style-type: none"> • Givet hur landskapet kopplat till vårddata utvecklas, samt det faktum att fler aktörer engagerar sig i området, behöver kontinuerlig omvärldsbevakning genomföras för att förstå hur andra aktörers arbete påverkar den regiongemensamma vårddatahubbens roll. • Rekommendationen behöver ta höjd för att ett antal strategiska initiativ kommer att realiseras på regional såväl som nationell nivå och kommer med stor sannolikhet att påverka vårddatahubbens arbete. • Vidare rekommenderar projektet att kontinuerligt bevaka hur regulatoriska förutsättningar samt dess riktlinjer (t.ex. EHDS) kommer påverka området för sekundäranvändning. 	<p>Ett antal stora och strategiska initiativ som berör hälsodataområdet genomförs för närvarande eller planeras i närtid att startas upp. För att undvika suboptimering kopplat till arbetssätt och resultat är en kontinuerlig omvärldsbevakning av såväl nationella och regionala initiativ mycket viktig.</p> <p>Omvärldsbevakningen kan användas som input såväl till ny kunskap som till inspiration.</p>	S9

Realisering av rekommendationerna ovan, föreslaget inom ramen för genomförande av:

1 Arbetspaket 1: Teknisk demonstration och kravställning

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Hänvisar till slutsatser från kapitel 4 i denna rapport.

Delprojektets rekommendationer (3/4)

#	Område	Föreslagen rekommendation och nästa steg	Syfte med framförd rekommendation	Slutsatser*
4.6	Utforma alternativ för paketering och realisering	<ul style="list-style-type: none"> Baserat på referensarkitekturen samt underliggande kravställning utveckla potentiella alternativ till slutgiltig paketering inklusive modeller för förvaltning och drift för jämförelse. Rekommendationen baseras på delprojektets genomförda dialog med ett antal centrala aktörer som potentiellt kan utgöra referensarkitekturens tekniska hemvist. Inom ramen för rekommendationen bör en översyn av existerande verktyg genomföras sett utifrån tekniskt, investerings- och avtalsmässigt perspektiv. Detta för att utvärdera om existerande komponenter kan återanvändas. 	<p>Rekommendationen föreslås att genomföras för att genomföra en inventering av vilka komponenter av referensarkitekturen som finns att tillgå via tillgängliga avtal och verktyg.</p> <p>Rekommendationen syftar vidare till att återge den strategiska inriktningen för hur datahubben kan tänkas att realiseras och etableras. Utformning av realiserings- samt paketeringsalternativ är en förutsättning för att på ett strukturerat sätt kunna genomföra kommande delsteg såsom marknadsdialog samt etableringsplanering.</p>	S8
4.7	Förberedelse inför samt genomförande av marknadsdialog	<ul style="list-style-type: none"> Baserat på paketeringsalternativ etablera en fördjupad förståelse för hur marknaden ser på framtagna referensarkitektur samt identifiera kostnadsdrivande aspekter och områden. Baserat på marknadsdialog utforma beslutsunderlag som underlag till potentiellt investeringsbeslut avseende tekniska komponenter. Aktiviteten genomförs förslagsvis enligt etablerad förstudiemetodik. 	För att kunna validera utvalda tekniska och investeringsmässiga aspekter av projektets referensarkitektur föreslås att en strukturerad marknadsdialog genomförs. Detta för att inhämta marknadens input på komplexa samt särskilt prisdrivande områden.	

Realisering av rekommendationerna ovan, föreslaget inom ramen för genomförande av:

2 Arbetspaket 2: Teknisk realisering

**Nationellt system
för kunskapsstyrning
Hälso- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

* Hänvisar till slutsatser från kapitel 4 i denna rapport.

Delprojektets rekommendationer (4/4)

#	Område	Föreslagen rekommendation och nästa steg	Syfte med framförd rekommendation	Slutsatser*
4.8	Utveckling av kostnadsmässig baslinje samt underliggande finansieringsmodell	<ul style="list-style-type: none"> Baserat på hur användandet av referensarkitekturen kommer att se ut behöver de underliggande mekanismerna och komponenterna utvecklas och tydliggöras. Rekommendationen tar sin utgångspunkt i såväl genomförd marknadsdialog som slutgiltigt valt paketeringsalternativ. För att på ett tydligt sätt kunna utvärdera och simulera framtida investerings- och kostnadsutveckling rekommenderas att en kostnadsmässig baslinje för nuvarande arbetssätt etableras och jämförs med information inhämtad ifrån marknadsdialog. 	<p>Givet den potentiella bredden av användande organisationer samt kostnadsmässiga påverkan, föreslås rekommendationen att genomföras baserat på vedertagen metodik för kostnadsestimering och nyttoanalys.</p> <p>För att få en helhetsbild behöver vi systematiskt relatera prisuppgifterna från marknadsstudien till resultaten av kostnadsestimeringen och nyttoanalysen.</p>	
4.9	Utforma etableringsplan	<ul style="list-style-type: none"> Baserat på delstegen ovan, utveckla en etableringsplan innehållande alternativ för drift- och förvaltning samt uppskattad budget. För att effektivisera planeringen och genomförandet bör arbetet utgå från den i projektet valda prioriteringen av utvalda användarfall. Inom ramen för rekommendationen föreslås även att en testmiljö etableras för ytterligare fördjupning av test av användningsfall. 	Framförd rekommendation syftar till att, i linje med framtida vald strategi för realisering, utforma en förankrad och genomförbar plan för den tekniska etableringen. Planen behöver ta höjd för regionernas förmåga att ansluta sig till vårddatahubben och vara avvägd i termer av insats, utfall och resultat.	S7
4.10	Datakonsekvensbedömning	<ul style="list-style-type: none"> Givet att den föreslagna referensarkitekturen till viss del samlar in data på ett nytt sätt och med hjälp av ett nytt arbetssätt rekommenderas att en risk- och datakonsekvensbedömning genomförs. Utöver aspekter kopplat till den inkluderade datan rekommenderas att även arbetssätt, verktyg samt rutiner inkluderas i denna bedömning. 	Rekommendationen syftar till att analysera och beakta de konsekvenser kopplat till datasäkerhet samt juridik som införande av vårddatahubben resulterar i.	S3

Realisering av rekommendationerna ovan, föreslaget inom ramen för genomförande av:

2 Arbetspaket 2: Teknisk realisering

**Nationellt system
för kunskapsstyrning
Hälsa- och sjukvård**

SVERIGES REGIONER I SAMVERKAN

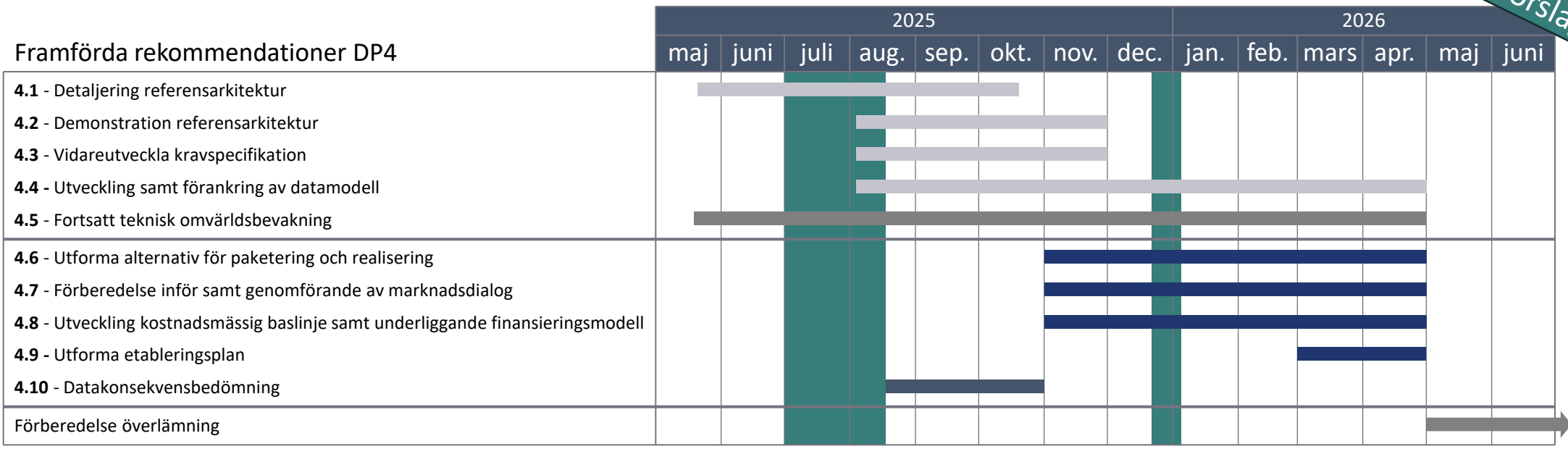
* Hänvisar till slutsatser från kapitel 4 i denna rapport.

Givet att framförda rekommendationer innehåller beroenden samt aktiviteter som behöver genomföras i sekvens, föreslås nedan struktur samt paketering för nästa steg

Förslag

AP1

AP2



Det är av yttersta vikt att respektive arbetspaket resurssätts samt bemannas med adekvat kompetens. Detta då de föreslagna aktiviteterna kommer att vara av teknisk komplex karaktär. Delprojektets aggregerade & uppskattade resursuppskattning (antal FTE*) framgår nedan

	Teknisk projektledare	Lösningsarkitekt	Data-modellering	Informatiker	Utvecklare	Analytiker	
Arbetspaket 1	1 FTE	1 FTE	1 FTE	0,5 FTE	1 FTE	1 FTE	5,5 FTE
Arbetspaket 2	1 FTE	1 FTE				2,5 FTE	4,5 FTE

Nationellt system
för kunskapsstyrning
Hälso- och sjukvård

SVERIGES REGIONER I SAMVERKAN

*FTE står för "Full-Time Equivalent" på engelska, vilket på svenska översätts till "motsvarande heltidsanställd"

Fördjupning Arbetspaket 1: Teknisk demonstration och kravställning

Övergripande syfte

Arbetspaketet syftar till att fortsatt utreda och detaljera delar av den föreslagna referensarkitekturen samt ingående komponenter i enlighet med föreslagen projektrekommendation. Inkluderat i arbetspaketet inkluderas även teknisk omvärldsbevakning.

Delar i arbetspaketet inkluderar aspekter kopplat till demonstration och utveckling av PoC i syfte att påvisa och validera tilltänkt funktionalitet.

Vidare inkluderas en aspekt kopplat till utveckling samt förankringsprocess avseende datamodell – ett område som inkluderar särskild teknisk och organisatorisk komplexitet.

Föreslagen leverans

- Leverans av demonstrationsmiljö [PoC] i validerings- och demonstrationssyfte
- Leverans av detaljerad kravspecifikation
- Detaljerad realiseringsarkitektur för identifierade, särskilt komplexa områden
- Utvecklad gemensam datamodell samt föreslagen förankringsprocess
- Genomförande av kontinuerlig teknisk omvärldsbevakning

Förslag

Föreslagen kompetens

- Teknisk projektledare
- Informatiker
- Lösningssarkitekt
- Analytiker
- Datamodellering
- Utvecklare

Framgångsfaktorer

- Projektgruppering med relevant teknisk kompetens
- Etablerade styrning- och förankringsprocesser
- Tydliga mandat och eskaleringsprocesser
- Tekniska förutsättningar för att på ett agilt sätt utveckla föreslagen PoC
- Etablerade kommunikationsvägar till kringliggande initiativ

4.1 - Detaljering referensarkitektur

Fördjupningsområden såsom:

- 4.1.1 – Mappning från regionala datamodeller
- 4.1.2 - Behörighetsstyrning
- 4.1.3 - Dataåtkomst
- 4.1.4 - Pseudonymisering

4.2 - Demonstration referensarkitektur

- 4.2.1 - Konceptuell arkitekturdesign
- 4.2.2 - Detaljering plan PoC
- 4.2.3 - Skapa PoC

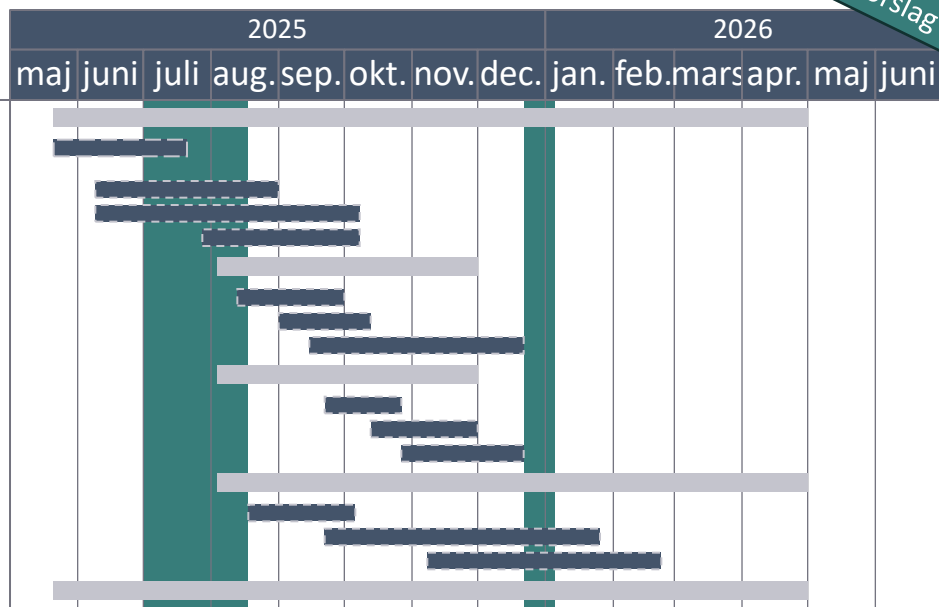
4.3 - Vidareutveckla kravspecifikation

- 4.3.1 – Detaljering kravspecifikation
- 4.3.2 – Förankring kravspecifikation
- 4.3.3 – Dokumentation kravspecifikation

4.4 - Utveckling samt förankring av datamodell

- 4.4.1 - Formering arbets- och förankringsprocess
- 4.4.2 – Framtagning av datamodell enligt förslagen arbetsprocess
- 4.4.3 - Detaljering ingående komponenter

4.5 – Fortsatt tekniskt omvärldsbevakning



AP2 Fördjupning Arbetspaket 2: Teknisk realisering

Övergripande syfte

Arbetspaketet syftar till att utreda och besluta om hur tekniska komponenter i vårddatahubben ska etableras samt dess organisatoriska och investeringsmässiga påverkan.

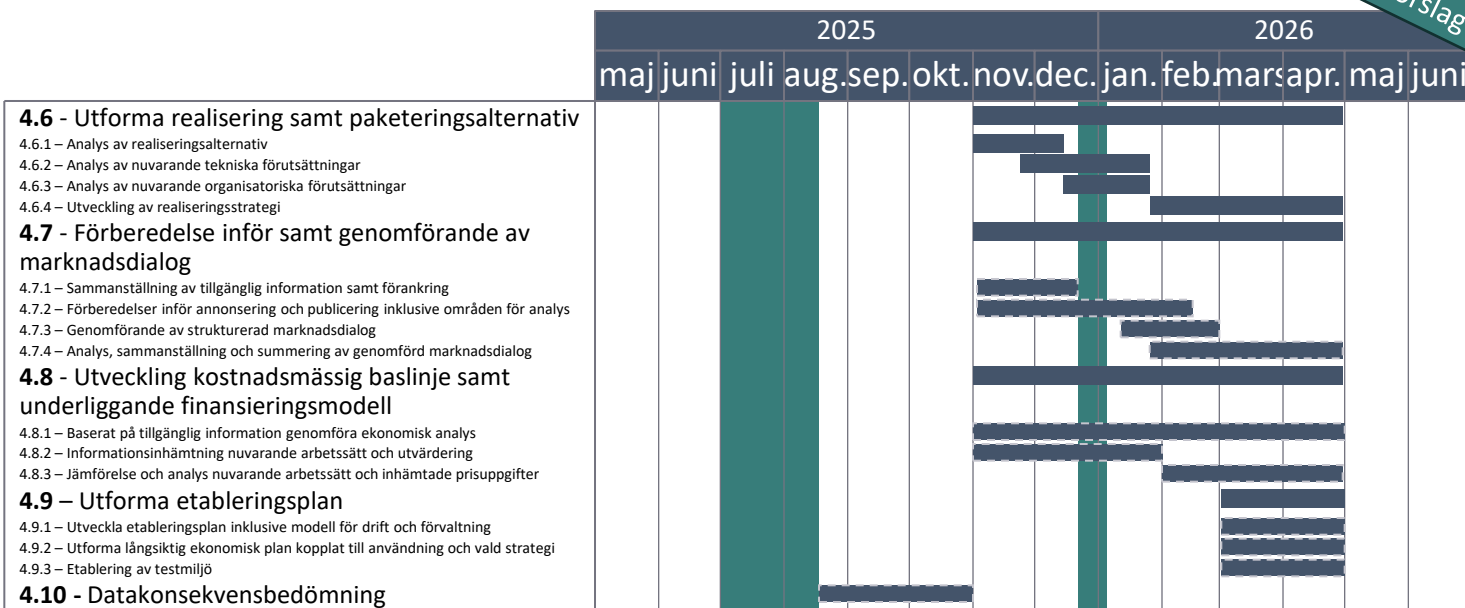
Delar i det föreslagna arbetspaket syftar till att utforma en övergripande strategi kring realisering och paketering, genomförande av marknadsdialog samt utveckla ekonomisk baslinje för användande av den framtida lösningen.

Inom ramen för aktiviteten inkluderas även att utforma en realiseringsplan samt genomföra en detaljerad datakonsekvensbedömning.

Föreslagen leverans

- Leverans av föreslagen strategi för realisering inklusive föreslagen organisatorisk hemvist
- Förberedelse samt genomförande av marknadsdialog för att inhämta underlag till investeringsbeslut samt marknadsinformation
- Upprättad ekonomisk baslinje samt underliggande finansieringsmodell för framtida användande
- Leverans av föreslagen etableringsplan

Förslag



Föreslagen kompetens

- Teknisk projektledare
- Lösningssarkitekt
- Datamodellering
- Informatiker
- Analytiker
- Informationssäkerhet skompetens

Framgångsfaktorer

- Projektgruppering med relevant teknisk och kommersiell kompetens
- Etablerade styrning- och förankringsprocesser
- Tydliga mandat och eskaleringsprocesser
- Förutsättningar för att på ett strukturerat sätt kunna inhämta och analysera marknadsinformation

**Nationellt system
för kunskapsstyrning
Hälso- och sjukvård**

Leverabel 4.4 Appendix (separata dokument)

- Appendix 1 Fördjupande bilder
- Appendix 2 Övergripande designkriterier
- Appendix 3 Regulatorisk översikt
- Appendix 4 Skillnader primär- och sekundäranvändning

This presentation was produced with the financial assistance of the European Union. Its content is the sole responsibility of the author(s). The views expressed herein can in no way be taken to reflect the official opinion of the European Union.

The project is funded by the European Union via the Technical Support Instrument, managed by the European Commission Directorate-General for Structural Reform Support.

This presentation has been delivered in May 2025, under the EC Contract No. 300089094



© European Union 2023

Unless otherwise noted the reuse of this presentation is authorised under the [CC BY 4.0](https://creativecommons.org/licenses/by/4.0/) license. For any use or reproduction of elements that are not owned by the EU, permission may need to be sought directly from the respective right holders.