

Hanseatic Life Science Research Infrastructure Consortium - Hanseatic Science Cloud

HALRIC | 2023

"Att förstå ekosystemet för en hållbar datakultur"

Emma-Lisa Hansson - Informationsspecialist & delprojektledare

LUNARC, Centrum för Tekniska och Vetenskapliga Beräkningar vid Lunds Universitet

Interreg



Co-funded by
the European Union

Öresund-Kattegat-Skagerrak



Om projekten

- **LU projektet**
 - Roadmap till ledning för fortsatt arbete med forskningsdata
- **Lagringsprojektet**
 - Initierades av Chalmers. Samarbete mellan digitala infrastrukturer (SND, SNIC, Sunet) och två universitet (Chalmers och Örebro univ)
- **Nationella nätverket**
 - Stort intresse och medverkan från svenska lärosäten och organisationer!
- **The Hanseatic Science Cloud (HSC)**
 - Molnlösningssprototyp för hantering och analys av stora och komplexa datamängder från forskningsinfrastrukturer. Ett interregfinansierat projekt 2023-2025

Från projektplanen

Lagringsprojektet (forskningsdata)

Syfte

- Föreslå en arkitektur
 - För forskares lagring av forskningsdata,
 - Lokala lösningar och gränssnitt gentemot externa lösningar.
- Kartlägga forskares behov av infrastrukturellt stöd för forskningsdatahantering
- Fokus på lagring forskningsdata.
- Bidra till ökad förståelse för forskningen och lärosätets behov
- Peka på områden som behöver utvecklas för att tillhandahålla ett optimalt stöd för forskningsdatahantering.

Effektmål

- Bild av infrastrukturellt stöd som behövs (gällande lagring av forskningsdata under olika delar av forskningsdatalivscykeln), inkluderande kopplingar till relevant stöd för t ex delning och publicering av data.
- Kan utgöra ett underlag för kravspecifikation för olika delar av det infrastrukturella stödet hos de nationella infrastrukturerna samt hos lärosätena,
 - inklusive lagringstjänster, överföringstjänster, och APIer.

Inom 2 år:

- Förutsättningarna att utveckla det infrastrukturella stöd för att hantera forskningsdata under hela datalivscykeln,
 - inklusive välfungerande gränssnitt mellan lokala, nationella och internationella aktörers e-infrastrukturlösningar.

Metoder

För verksamhetsprojekten

- Compliance-lista: utveckling av grundlistan från LU-arbetet, nya visualiseringar.
- Intervjuer med forskare
- Visionssamtal med stödorganisationer och -funktioner
- Intervjuer stödorganisation
- Månatliga avstämningar med LU-rektorer
- Avstämningar projektgruppen
- User stories

Intervju

Lagringsprojektet

- Individuella intervjuer med forskare
 - Semistrukturerad kvalitativ intervjumetod
- 21 stycken intervjupersoner från olika forskningsfält
- Skapa en "heltäckande" bild av forskning i Sverige
- Forskare verksamma vid Chalmers tekniska högskola och Örebro universitet
- Digitala videointervjuer som transkriberades
- Två projektmedlemmar genomförde varje intervju

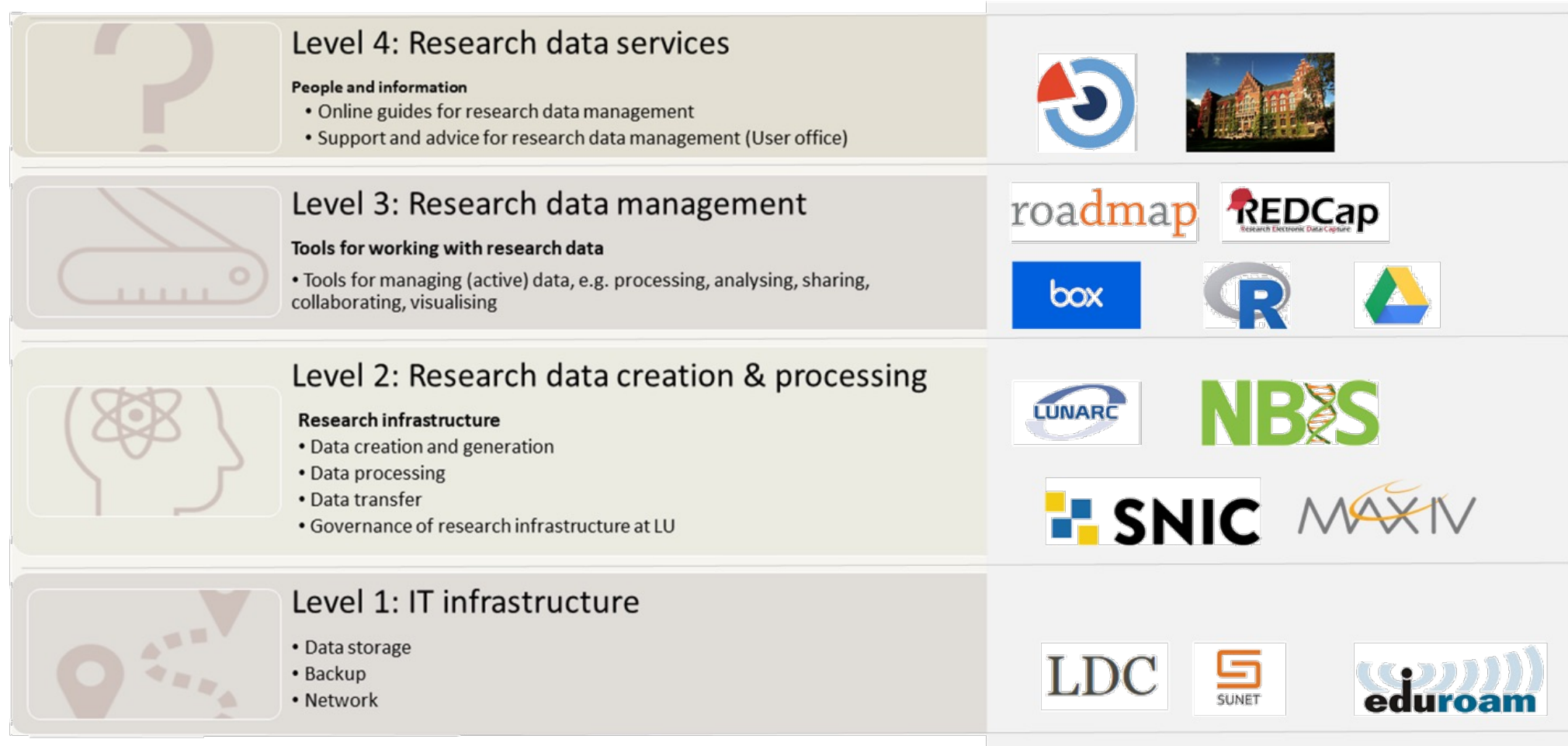
Intervjumall

Lagringsprojektet

- Kort introduktion till projektet och kartan
 - Utgår från ett projekt primärt
- Kort om forskningen som intervjupersonen genomför
- Forskare får fritt beskriva dess hantering av forskningsdata under de olika faserna i kartan
 - Stödfrågor till varje fas vid behov
 - Kompletterar med nöjdhet kring hur varje fas hur det fungerar idag
 - Grön: nöjd
 - Gul: ok men det finns förbättringar som måste göras
 - Röd: missnöjd
- Avslutar med drömscenario

Underlag intervjuer

Nivåer i infrastrukturen



En nationell karta

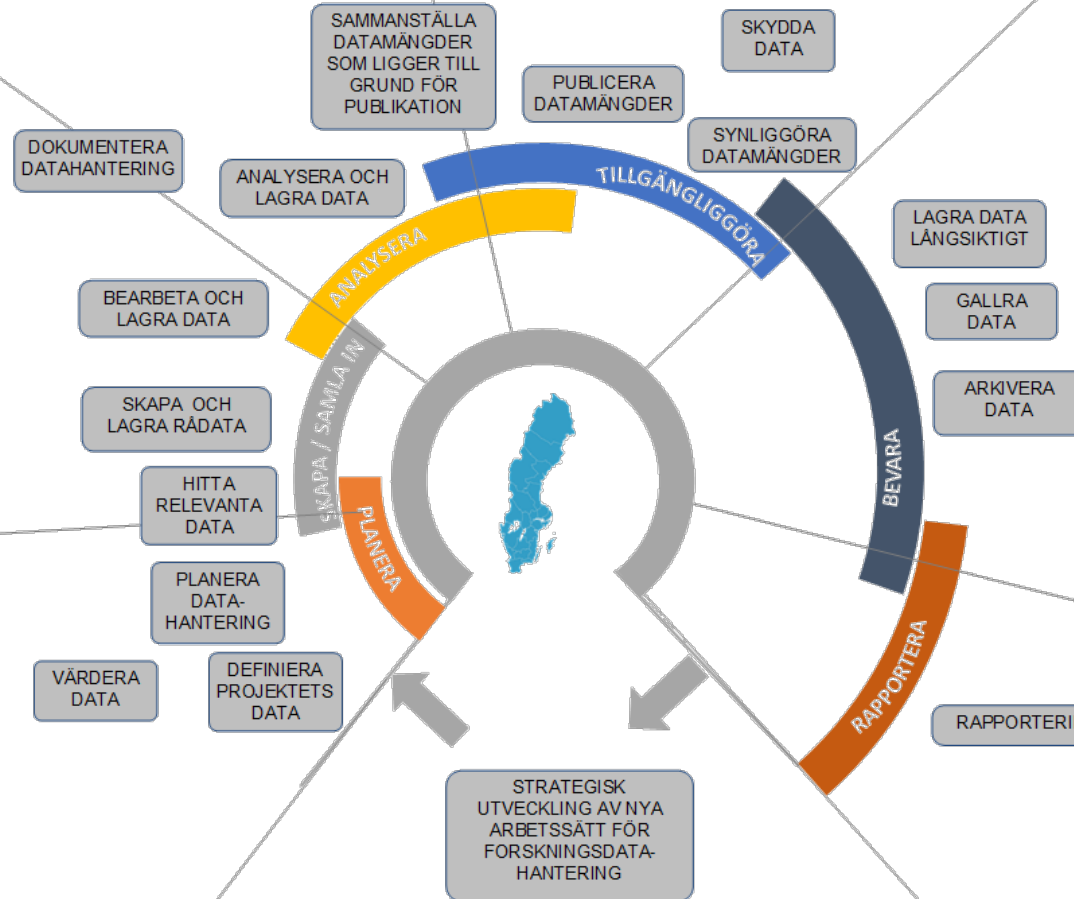
över forskningsdataområdet

- Vilka förmågor behöver Sverige ha för att hantera forskningsdata?
- Hur ser nuläget ut?
- Vad är idealläget?
- Hur når vi till idealläget?

Kartornas funktioner

- Pedagogisk modell: hjälp att förklara ett komplext område
- Gemensamt (bild)språk
- Ankarmodell: alla kan känna igen sig i kartan
- Konkretisering av krav, villkor, lagar, styrdokument etc

- Lärosäten och e-infrastrukturer ritar kartan tillsammans

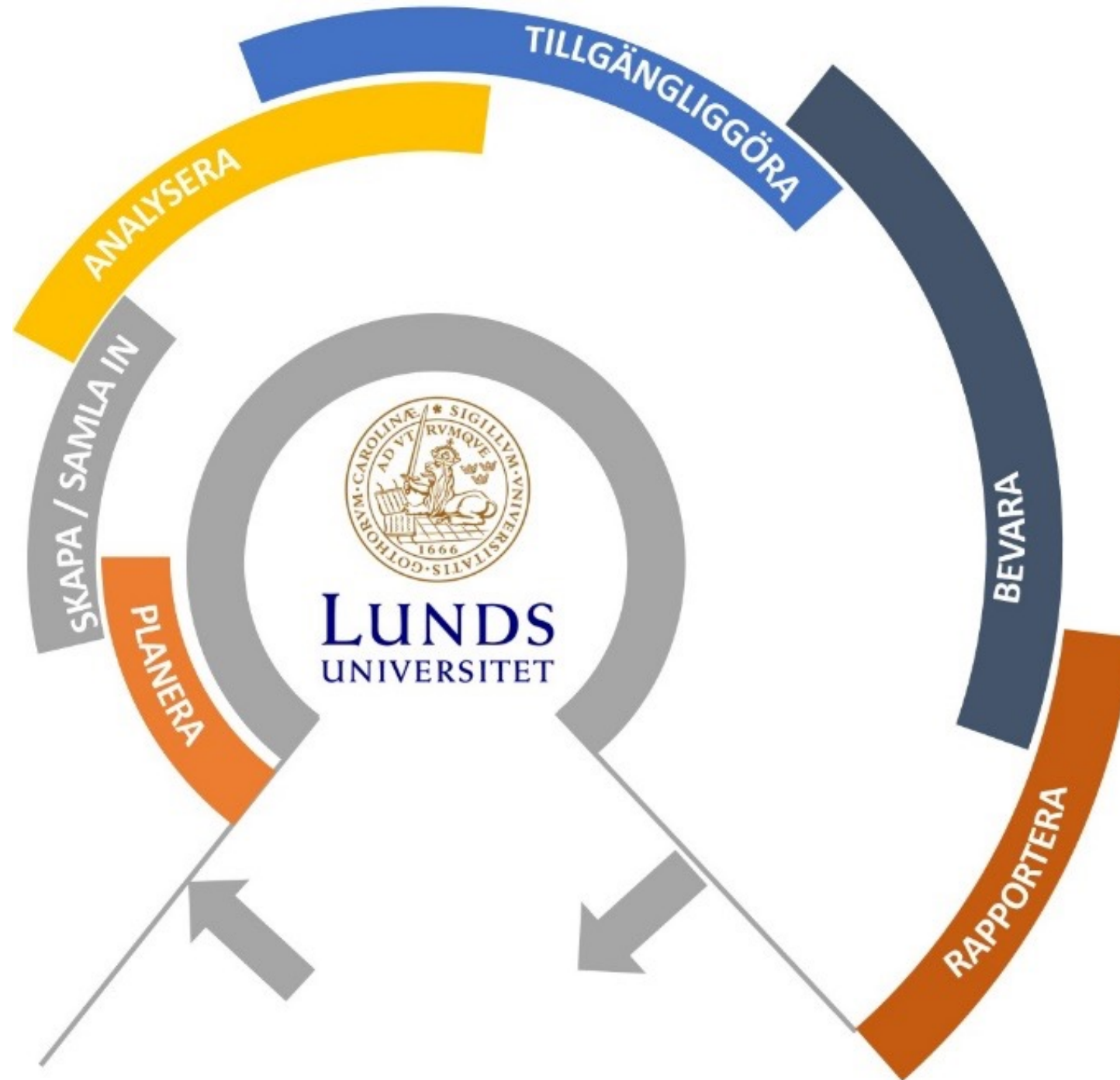


Huvudstegen

i kartan

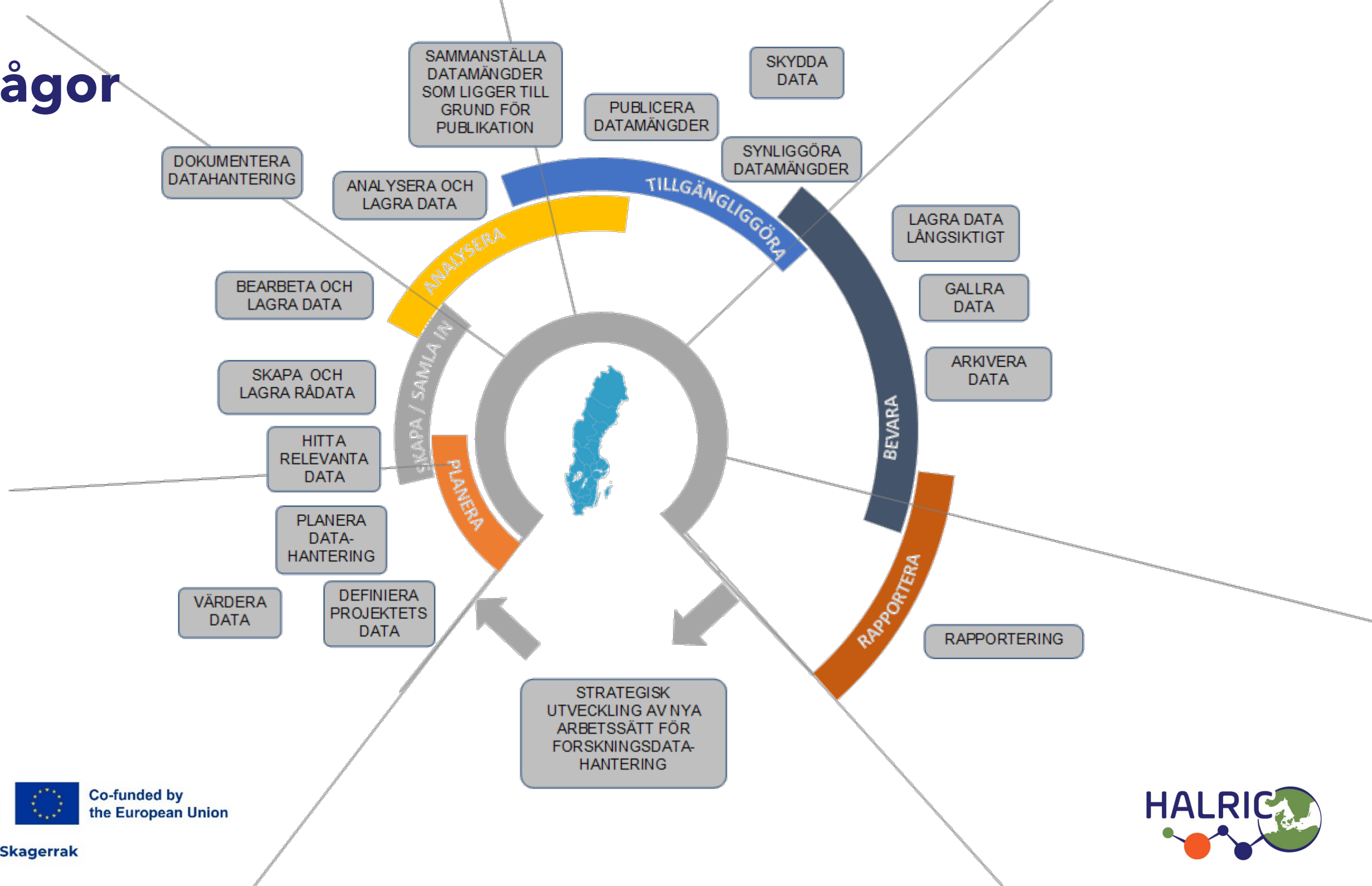
Huvudstegen i kartan är

- Planera
- Skapa/Samla in
- Analysera
- Tillgängliggöra
- Bevara
- Rapportera



Förmågor

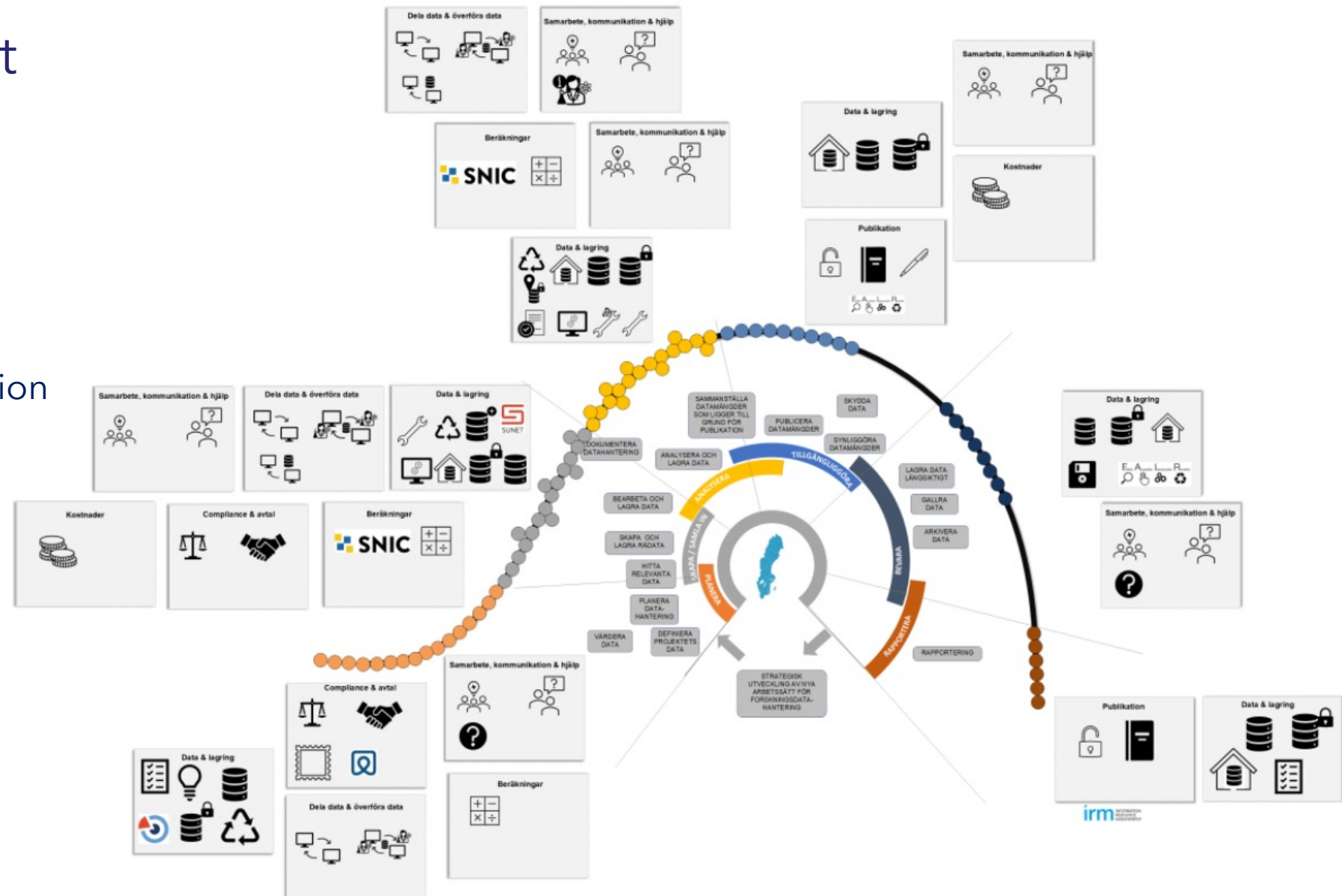
I kartan



Nulägesbeskrivning forskare

Lagringsprojektet

- Publikation
- Dela & överföra data
- Data & lagring
- Compliance & avtal
- Beräkning
- Samarbete, kommunikation & specialiserat stöd

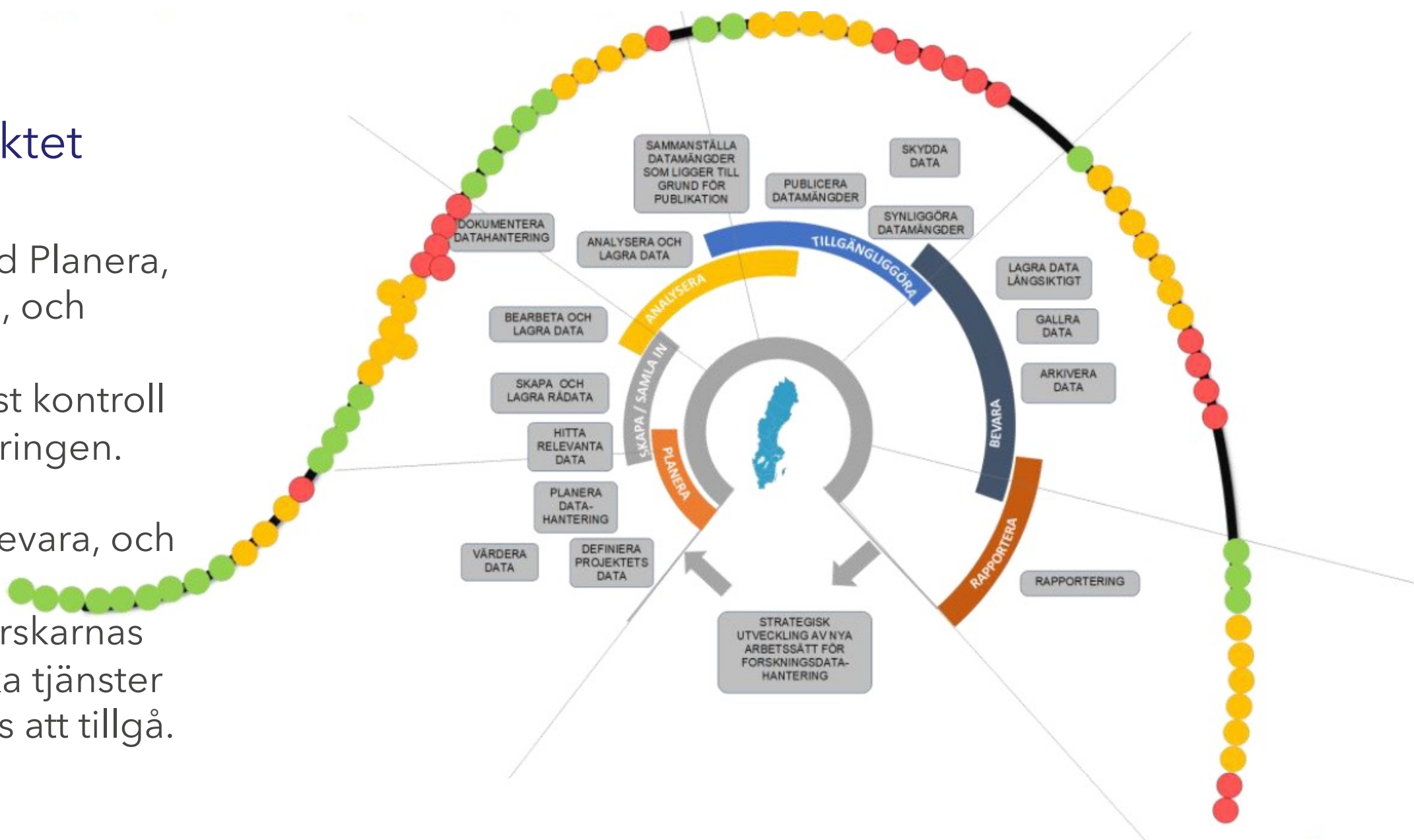


Nöjdhet

Lagringsprojektet

- Mest nöjda med Planera, Skapa/Samla in, och Analysera.
- Faser med störst kontroll över datahanteringen.


Tillgängliggöra, Bevara, och Rapportera, lägre mognadsgrad i forskarnas praktiker och i vilka tjänster och stöd som finns att tillgå.



Förmåge-exempel

Sektor tillgängliggöra, förmåga skydda data (LU projektet)

Förmåga



Ansvarig: Rektor

Personer:

Uppgift:

- Data hanteras säkert under den tid forskningsprojektet pågår och skyddas också i långtidsbevaring, dock uppfyller dessa tjänster inte några av LU beslutade risk- och säkerhetsnivåer.

Erbjuder andra förmågor:

- Alla förmågor berörs av skydda data.

Stödjer touch points i forskarresan:

- Skydda data finns med i hela processen.

Tillgängliggöra - Skydda data

Nuvarande status

Bra

- Ledning är ytterst ansvarig för att metoder och lagar följs för "skydda data".

Mindre bra

Inte bra alls

- Forskarnas behov av bättre lösningar och riktlinjer kring känsliga data är stora. Forskarna ställs inför stora utmaningar vad gäller känslig data under hela datalivscykeln.
- Forskarna saknar en yta för aktiva och inaktiva data där de kan dela, lagra och samarbeta på ett säkert sätt.
- Avseende skydd vid hantering, spridning, insamling av data/information är stödet undermåligt och strider därför även mot lagkrav från tex GDPR, etiklagen och andra lagrum (+23).

IT-stöd:

- LUSEC plattform för att lagra, hantera och analysera data på ett säkert sätt och i enlighet med nya GDPR.
- LUNARC scientific and technical computing at Lund University.
- IT-stöd erbjuds via LDC.

Dessa tjänster är inte bedömda eller uppfyller av LU beslutade risk och säkerhetsnivåer då dessa saknas.

irm INFORMATION RESOURCE MANAGEMENT

Förmåga



Onskad status:

För att uppnå rätt säkerhet och skydd av forskningsdata inom universitetets verksamhet ska forskare i framtiden kunna få stöd i ett systematiskt och riskbaserat informationssäkerhetsarbete. Referens till ett ledningssystem för informationssäkerhet på LU-nivå (LIS).

IT-stöd status:

- Målbild av IT-stöd formuleras utefter kravbild för ledningssystem för informationssäkerhet.

Vad pågår just nu:

- CISO (LUs Chief information security officer) leder och rapporterar arbetet till rektors ledningsråd avseende införandet av Ledningssystem för informationssäkerhet på LU-nivå (LIS) Dnr STYR 2021/1129: <https://www.medarbetarwebben.lu.se/artikel/nytt-ledningssystem-informationssakerhet-lis> Arbetet beräknas till ca 2-3 år.

Vad finns planerat i nästa steg (slutmålet):

- Se handlingsplan rektorsbeslut informationssäkerhet Dnr STYR 2021/1129.

Vad ser ni framåt?:

- Skydda data är ett integrerat arbetssätt som genomsyrar hela verksamheten.
- Forskare kan arbeta med forskningsdata på ett säkert sätt.
- Arbetsätt finns framtagna gällande konfidentialitet, integritet, tillgänglighet, spårbarhet och mottagande av data.

Tillgängliggöra - Skydda data

Förväntade utvecklingssteg:
Skydda data måste ses i olika perspektiv och med tydlig referens till LIS. Arbetssättet måste vara riskbaserat och systematiskt och innefattar bland annat:

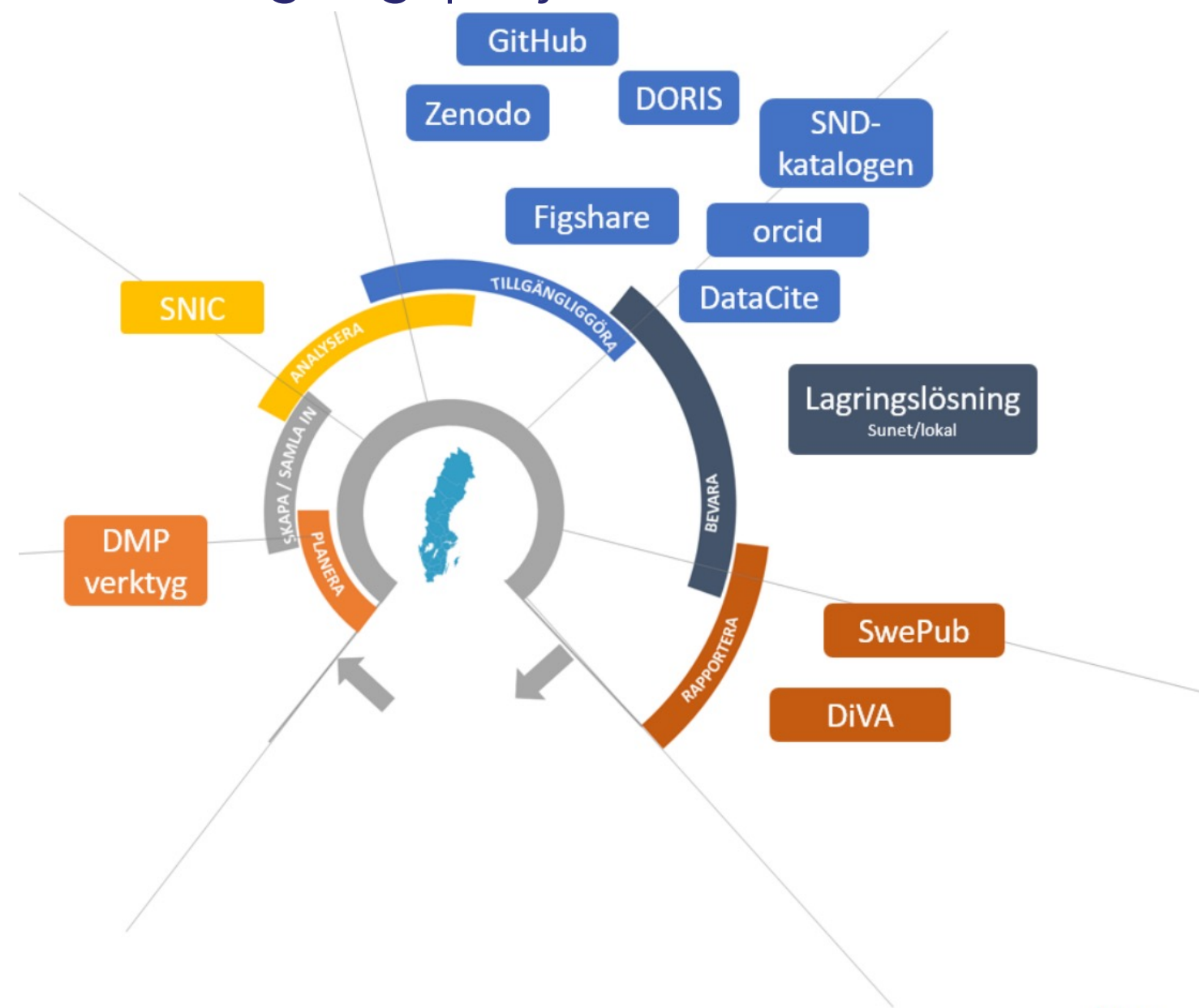
- Mål: Ägare av process och information (identifiera skyddsvärd information och persondata, klassificera utifrån konsekvens och skada).
- Mål: Ägare och förvaltare (var finns skyddsvärd information/data: skapas, lagras, hanteras, eller kommuniceras. Bärare av information: byggnader, processer, IT-system, IT-infrastrukturer, leverantörer, "cloud", personer, Internet).
- Mål: Förvaltas av ägare och förvaltare (utvärdera regelbundet. Vad & krav, baseras på MSBFS, ISO2700002, klassificering, typ av data, lagar och mer. GAP-analys, vilket skydd finns på plats, riskanalys).
- Mål: Säkerhetsåtgärder (införa/förändra skydd och reducera risken till rätt nivå för LU).

irm INFORMATION RESOURCE MANAGEMENT

Vanligt förekommande system och verktyg

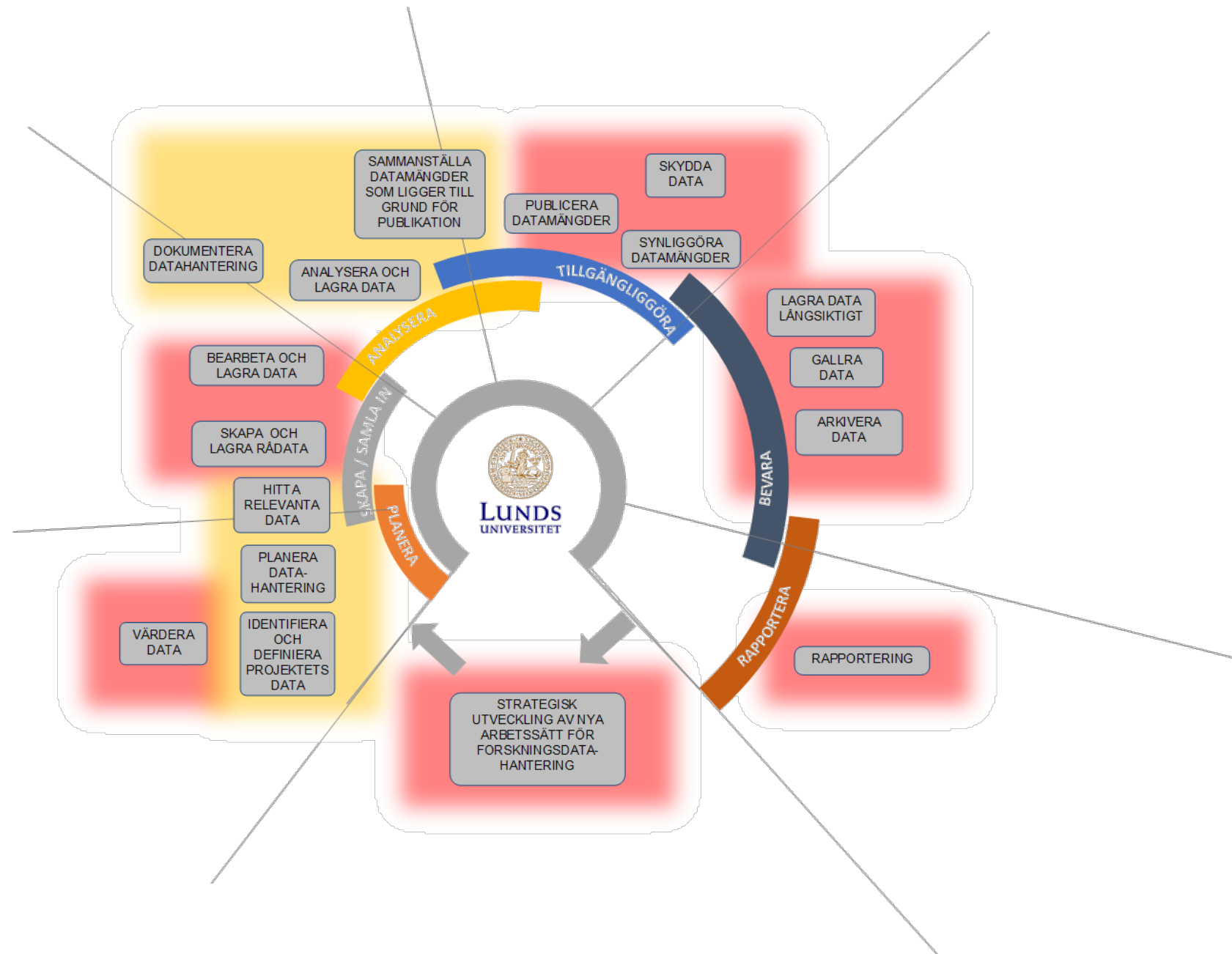
Som forskare interagerar med - lagringsprojektet

- Många verktyg för att hantera tillgängliggörande, viss integration mellan dessa.
- Lagringslösningen behöver bättre integration med övriga system



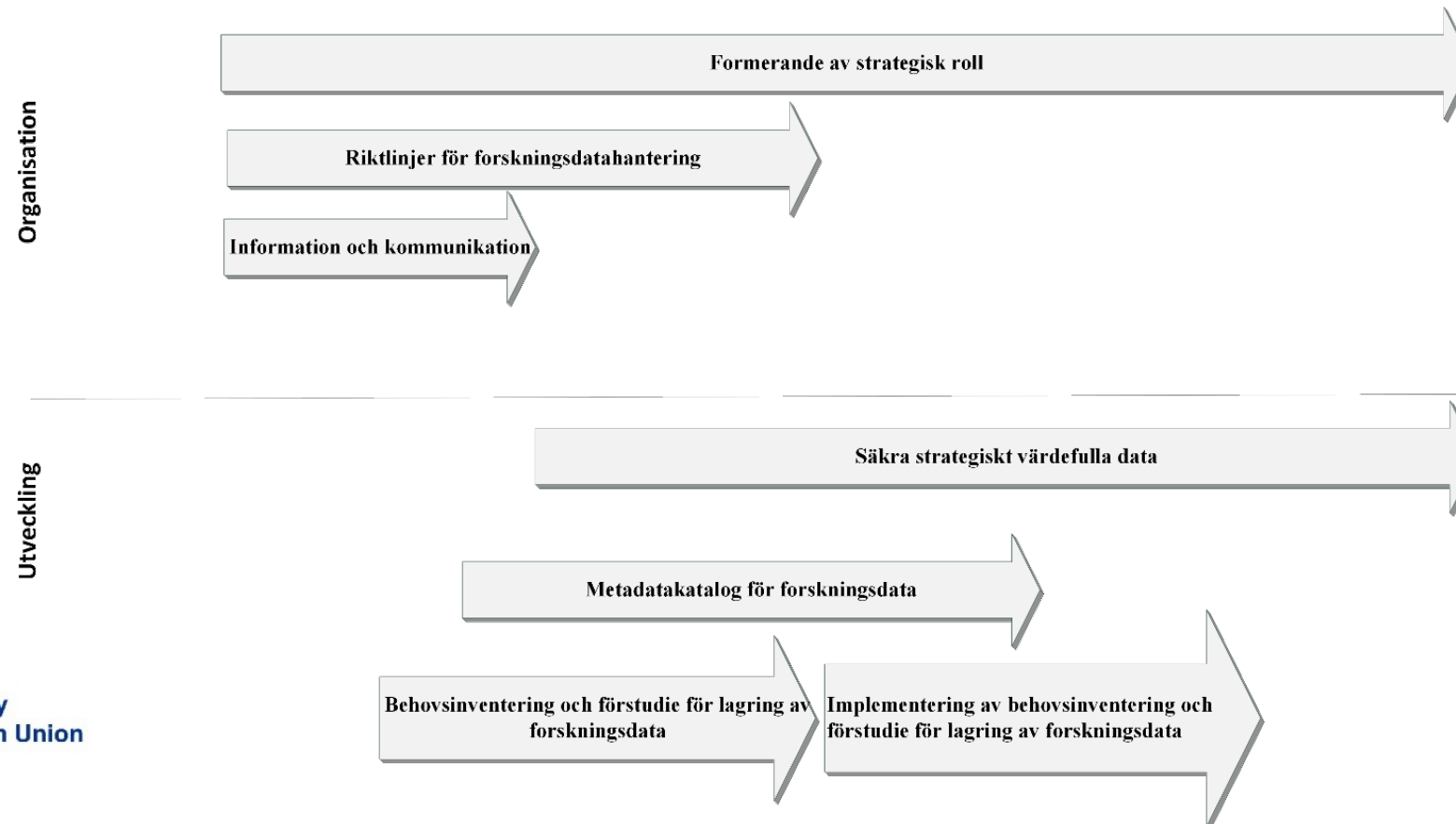
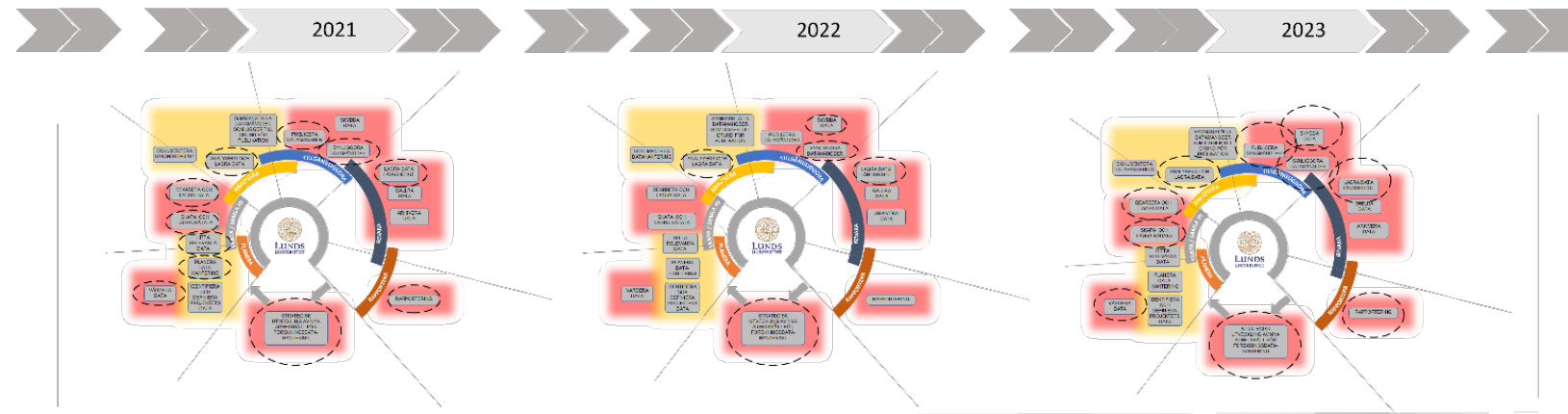
Heatmap

Lu-projektet



Roadmap - förslag på en färdplan

LU projektet



Designförslag

Steg för steg mot forskarnas drömscenario -User stories, eller användarhistorier

Sektor Analysera

Förmåga Dokumentera datahantering

1. Som användare av SNIC:s beräkningsresurser vill jag använda Jupyter Notebooks för att dokumentera min forskningsprocess.
2. Som forskare med samarbeten med industrin vill jag ha en "leased line"/pipeline/dedicated line för att säkerställa bandbredd och säkerhet i överföring av data.

Förmåga Analysera och lagra data

1. Som forskare vill jag snabbt kunna få tillgång till att analysera relevanta data utan att behöva ladda ner hela datamängder för att effektivisera min forskning.



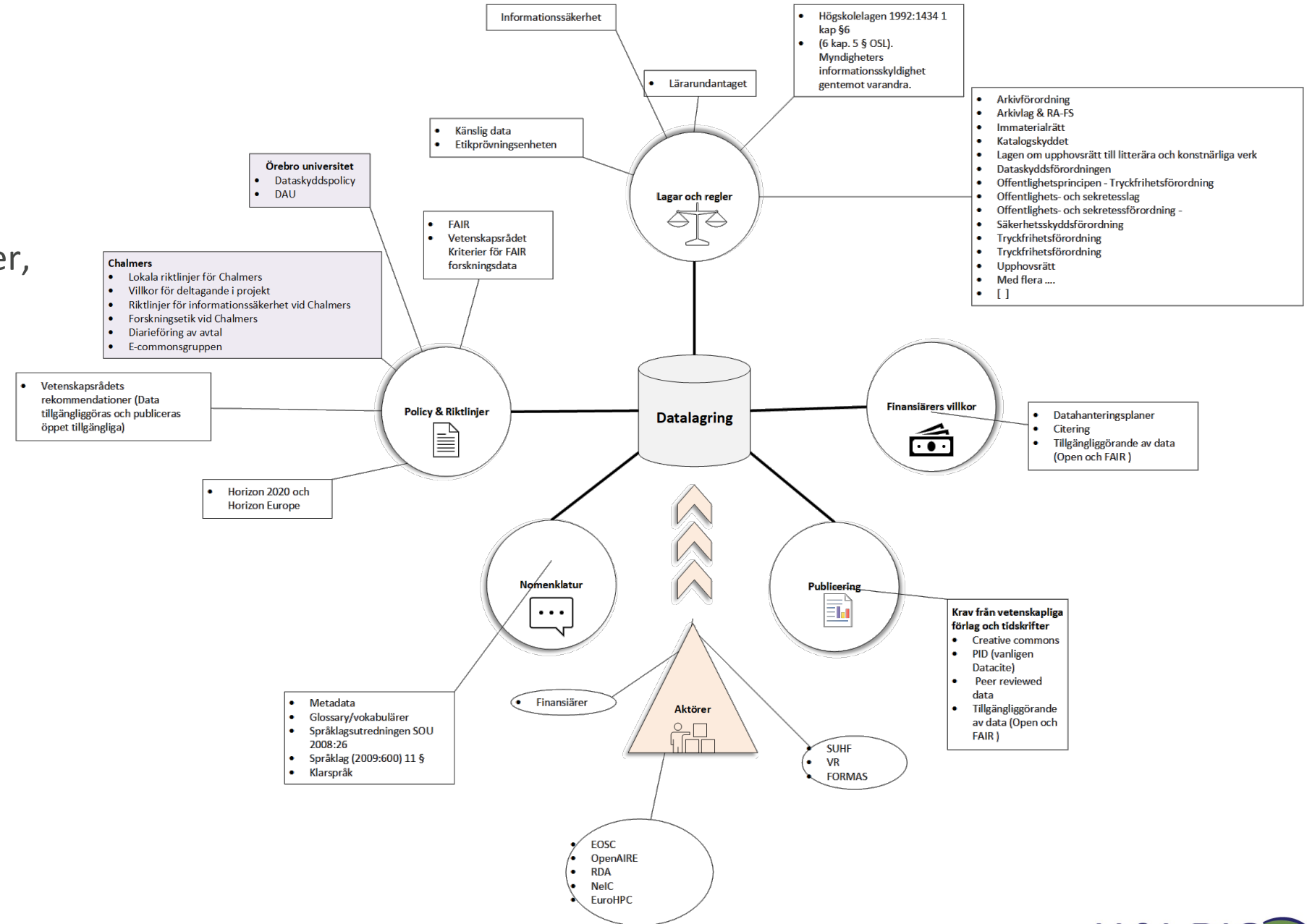
Utdrag ur
lagringsrapporten

Compliance-lista

Lagringsprojektet

Tar upp många aspekter, intressenter, info till heatmap, ämnen, internationellt/nationellt, nivå på minimumnivå

- Lagar & regler
- Finansiärers villkor
- Publicering
- Policy & riktlinjer
- Nomenklatur
- Aktörer



Sammanfattning av upplevelse intervju - lagringsprojektet

Lagringsprojektet

- Forskarnas återkoppling av intervjun - lätt att följa kartans faser
- Gavs möjlighet att revidera eller tydliggöra deras uttalande i samband med intervjun, då de såg övergripande anteckningar för varje fas
- Gav möjlighet för forskaren att verbalisera sin datahantering
 - Viktigt att ta vara på för fortsatta arbetet, från tyst kunskap till konkret kunskap
- Analyserade transkriberingarna i Nvivo



HALRIC:s vision är att öka innovationskapaciteten inom Öresund-Kattegat-Skagerrak-Hamburg Life Science-sektorn, genom ökad tillgång till forskningsinfrastrukturer och användning av gränsöverskridande front-end-teknologier, instrument, expertis och datahanteringslösningar.

Interreg

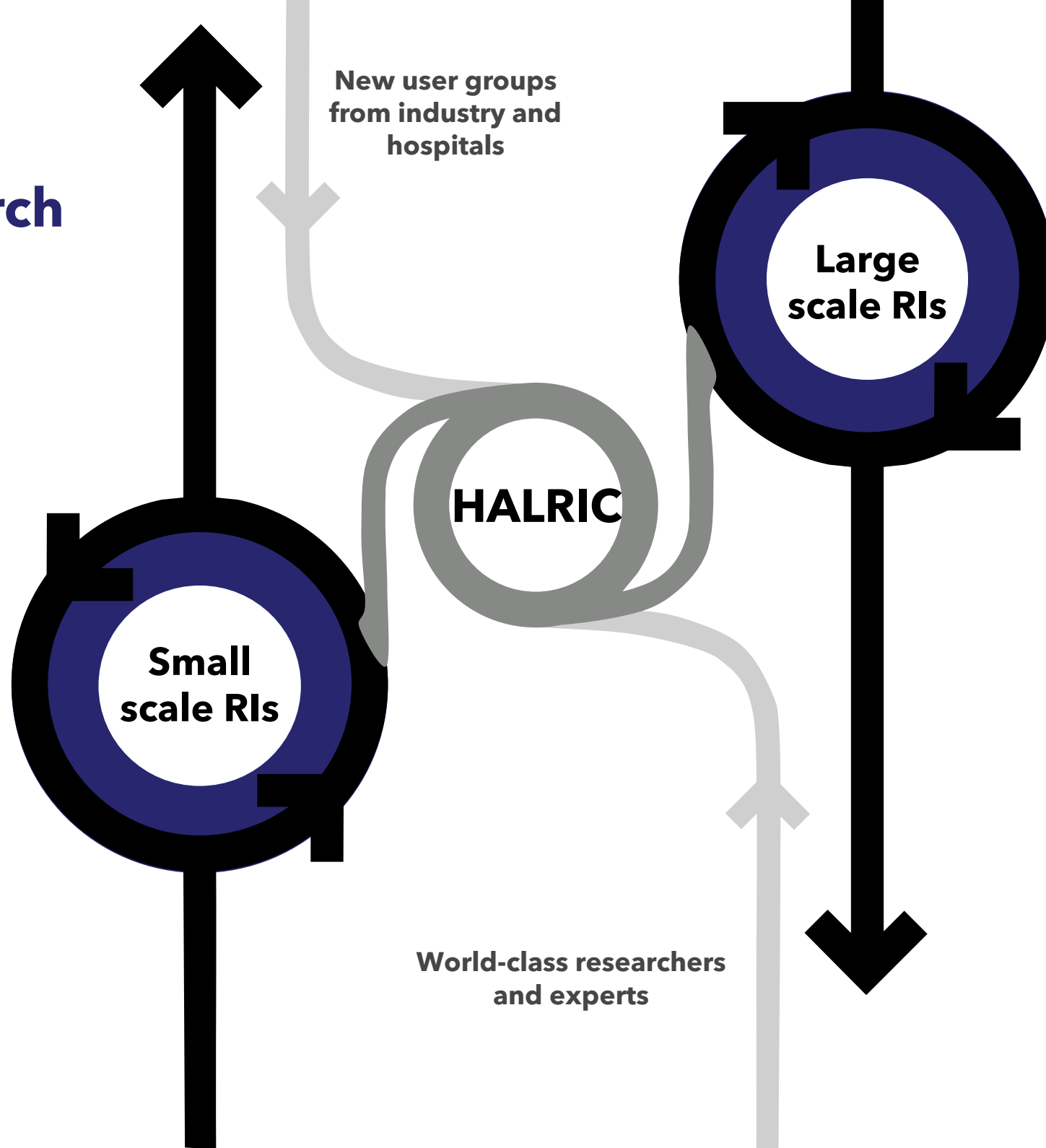


Co-funded by
the European Union

Öresund-Kattegat-Skagerrak



Increased **utilization** of globally **unique Research Infrastructures**



The Hanseatic Science Cloud (HSC)

A uniform entry point to data and compute resources

- HSC kommer att utveckla molnlösningssprototyp för hantering och analys av stora och komplexa datamängder från forskningsinfrastrukturer.
- HSC kommer att bidra till att innovationspotentialen i regionen ÖKS-Hamburg genom att förbättra gränsöverskridande tillgänglighet till forskningsinfrastrukturer.
- Ovanstående, i kombination med samarbetet kring data erbjuder HSC potentialen att möta FN:s hållbara utvecklingsmål.



Interreg



Co-funded by
the European Union

Öresund-Kattegat-Skagerrak





21 PARTNERS Hospitals, Universities, Research Infrastructures, Regions and Life Science Clusters



Interreg



Co-funded by
the European Union

Öresund-Kattegat-Skagerrak



Storskaliga forskningsinfrastrukturer (exempel)



Petra III, DESY Synchrotron
Hamburg, Germany



European XFEL
Schenefeld, Germany



ASTRID2 Aarhus University
Aarhus, Denmark

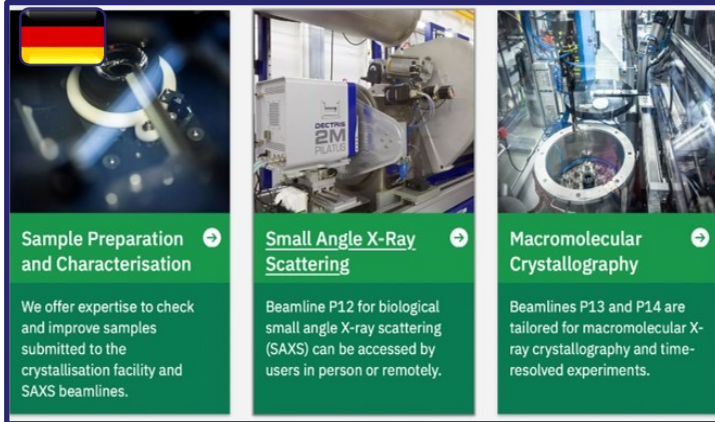


MAX IV, Lund University
Lund, Sweden



ESS – European Spallation Source
Lund, Sweden

Kompletterande forskningsinfrastrukturer (exempel)



Sample Preparation and Characterisation
We offer expertise to check and improve samples submitted to the crystallisation facility and SAXS beamlines.

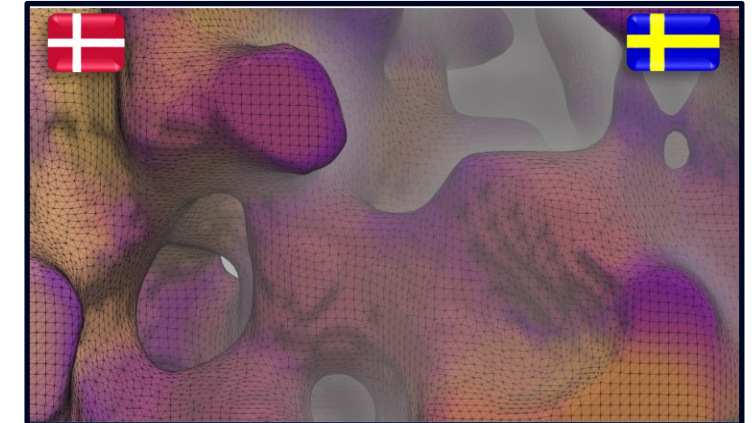
Small Angle X-Ray Scattering
Beamline P12 for biological small angle X-ray scattering (SAXS) can be accessed by users in person or remotely.

Macromolecular Crystallography
Beamlines P13 and P14 are tailored for macromolecular X-ray crystallography and time-resolved experiments.

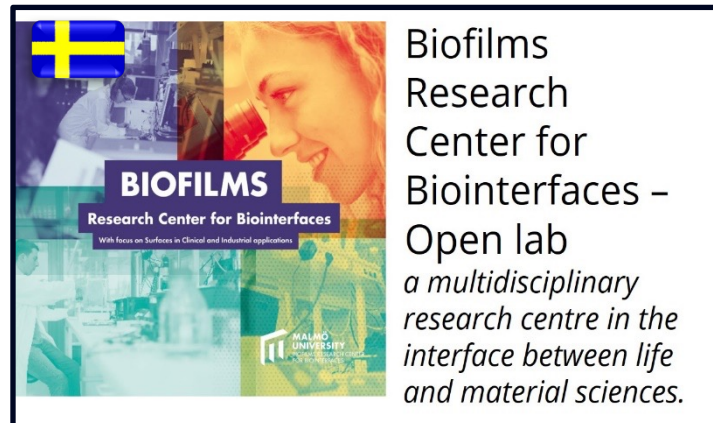
EMBL  **EMBL**
Hamburg, Germany



Life Science Building at Oslo University
Oslo, Norway



 **Center for Quantification of Imaging Data from MAX IV**
Copenhagen, Denmark & Lund, Sweden



BIOFILMS
Research Center for Biointerfaces
With focus on Surfaces in Clinical and Industrial applications

Biofilms Research Center for Biointerfaces – Open lab
a multidisciplinary research centre in the interface between life and material sciences.

 **Malmö University**
Malmö, Sweden



BMC: LBIC & SciLifeLab Site at LU
Lund, Sweden

Förmågor

För en akademisk molntjänst

- Användarnas behov och tjänsteleverantörernas förmågor utreds och matchas
- De kartlagda tjänsterna behöver inte finnas hos samarbetspartnerna utan kan vara delar av projekt och insatser som görs hos andra organisationer.
- HSC kommer att möjliggöra gränsöverskridande användning av beräknings- och lagringsresurser och säkerställa att de lagliga kraven och de nationella riktlinjerna uppfylls.
- Lagkraven och nationella policyer kan sätta begränsningar för hur data kan överföras över gränserna, hur beräkningsresurser kan delas och följaktligen vilka resurser som data behöver besökas av systemet snarare än överföras.
- Level of assurance (LoA).

Sammanfattning av resultat och reflektioner

Lagringsprojektet och HSC

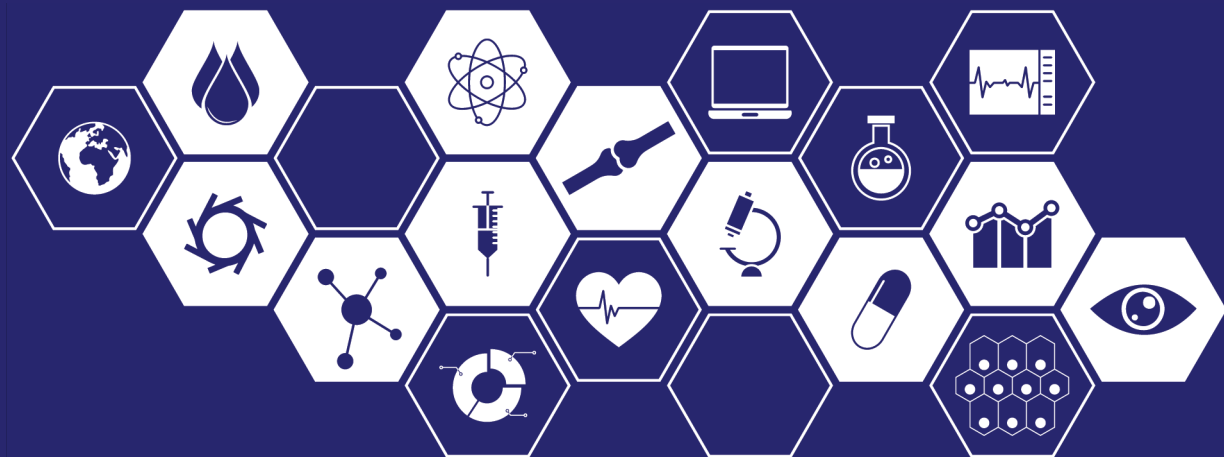
- Värdet av data är mycket högt, det är alla överens om. Men:
- Forskarnas kunskap om datahantering är tyst – flera berättade att de aldrig verbaliserat det som intervjuerna tog upp.
- Nomenklaturen måste överbryggas mellan forskning och stöd för att komma vidare. T.ex. "Vad är forskningsdata?"
- Juridiska frågor är utpekade som särskilt svåra och särskilt viktiga för forskare och stödfunktioner: Hur göra rätt? Vilken tolkning stämmer? Bl.a. för samarbeten och att dela data med andra.
- Forskningen är global; nationella och lokala ramverk sätter upp artificiella gränser för forskningsdatahanteringen. Workarounds.
- Incitament för publicering av data saknas för många. Verktyg för analys kan ge mycket credit i forskningsområden, men publicering av data är långt borta.

Ringar på vattnet...

Det nationella projektets resultat används av och byggs vidare på av SND, Chalmers, Örebro och KTH.

Förmågor och verksamhetsarkitektur har kommit in i "vokabulären" hos Vetenskapsrådet och SUHF (Sveriges universitets- och högskoleförbund)

LU arbetar med metadata för forskningsdata i sitt CRIS-system samt infört en projektledarroll för e-infrastrukturutveckling.



Tack!
Frågor?

Interreg



Co-funded by
the European Union

Öresund-Kattegat-Skagerrak



www.halric.eu

Tack! Håll gärna kontakten!



Kristina Jacobsson

Verksamhetsarkitekt och verksamhetsutvecklare på E.ON Energidistribution

[LinkedIn:](#)

[linkedin.com/in/kristina-jacobsson/](https://www.linkedin.com/in/kristina-jacobsson/)



Emma-Lisa Hansson

Research data architect på Lunds universitet

[LinkedIn:](#)

[linkedin.com/in/emma-lisa-hansson-928365201/](https://www.linkedin.com/in/emma-lisa-hansson-928365201/)



Cecilia Nordén

Verksamhetsarkitekt på IRM

[LinkedIn:](#)

[linkedin.com/in/cecilianorden/](https://www.linkedin.com/in/cecilianorden/)