



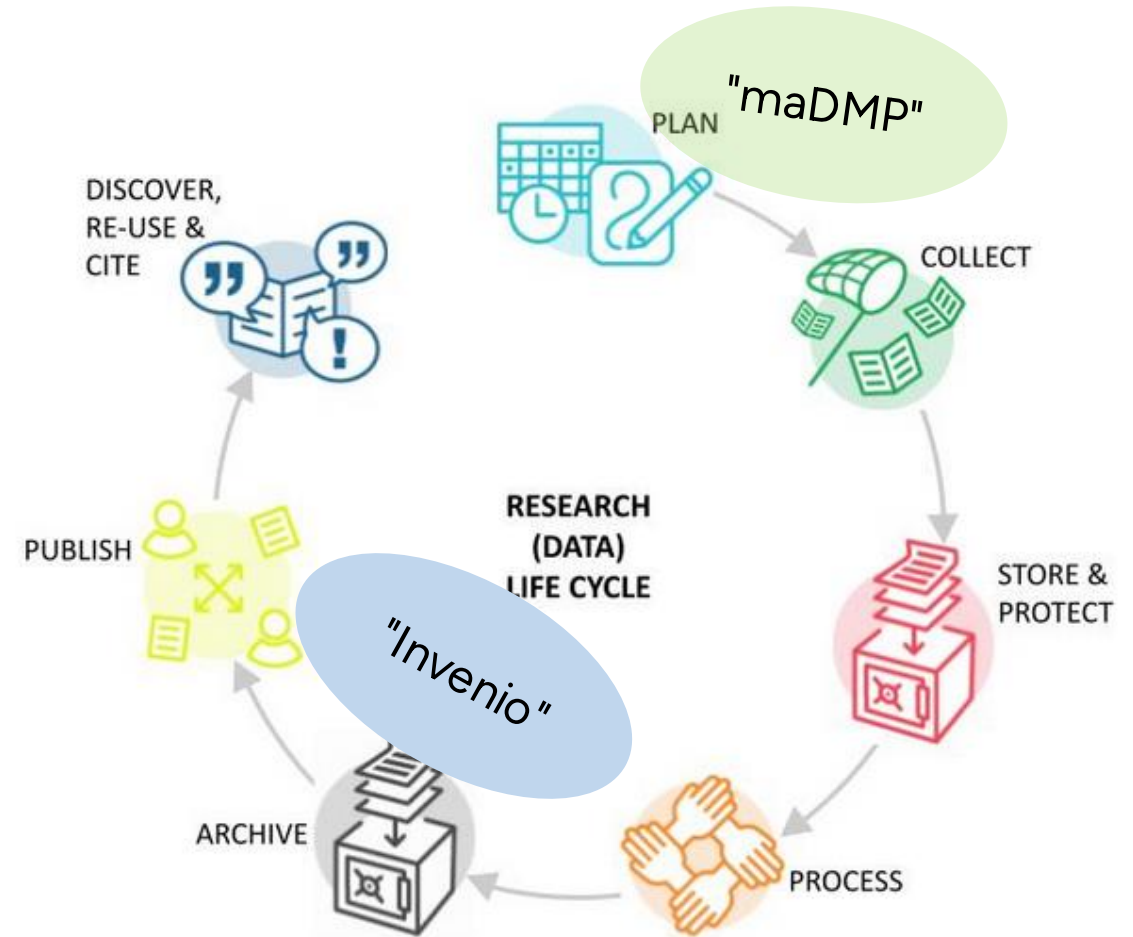
Införandeprojekt KTH Data repository

Ett utvecklingsprojekt inom KTH:s delportfölj för forskning

Övergripande mål: att stötta forskares förmåga att bedriva forskning genom att förbättra tillgång till data av hög kvalitet i alla steg i forskningsprocessen

Tidigare och pågående projekt:

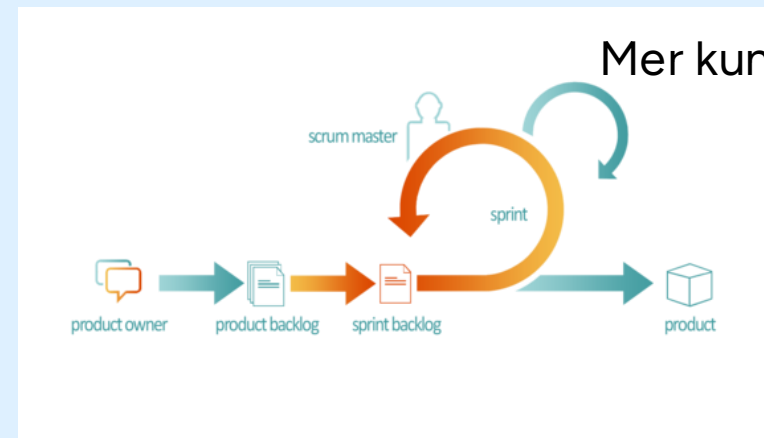
- Införande: verktyg datahanteringsplan
- Behovsinsamling "forskar-resan"
- Behovsinsamling "data-intensiv forskningsmiljö"
- Benchmarkinganalys –hur ser stöd ut på andra universitet?
- Maskinläsbara Datahanteringsplaner: för minskad manuell administration
- **Invenio- digitalt repositorium: För att stötta förmågorna beskriva, arkivera, publicera och återanvända data inom forskningsprojekt**



Hur har vi jobbat?

- "Agil-ish", användarcentrerat
- Tvärfunktionellt team sen 2019 (KTHB, RSO, IT, PDC/EECS, ca 3 FTE)
- I Invenio-projektet: KTHB /IT/ konsulter 1.5-2.5 FTE (+CERN/open source partners)
- Ett större antal demos för 54 unika testanvändare
- Input från test-användare -> vidareutveckling
- Löpande support och rådgivning -> analys: vad frågas det mest om? Vilket stöd saknas/är bristfälligt
- Workshops för doktorander & forskare -> frågor -> analys

Vad har vi lärt oss?



- Olika behov i olika forskningsdomäner och för olika forskningsmetodik
- Stora behov kring delning av data i samverkan
- Utmaningar kring infoklassning – konfidentialitet ändras under forskningsprocess på olika sätt i olika domäner
- Vi ser återkommande mönster där vi kan se use cases med liknande behov

Invenio-projektet: initial kravbild och status nuläge

Som forskare vill jag kunna:

- Ladda upp dataset och beskriva dem
- Tilldela begränsad åtkomst till data
- Söka efter mina och andras data
- Begära åtkomst till data
- Ha åtkomst till data via API
- Ha åtkomst i realtid och över längre tid, även efter att forskningsprojekt är slut
- Samla och analysera olika versioner av data för användning i forskningsresultat, ansökan eller publikation



Som forskare vill jag kunna:

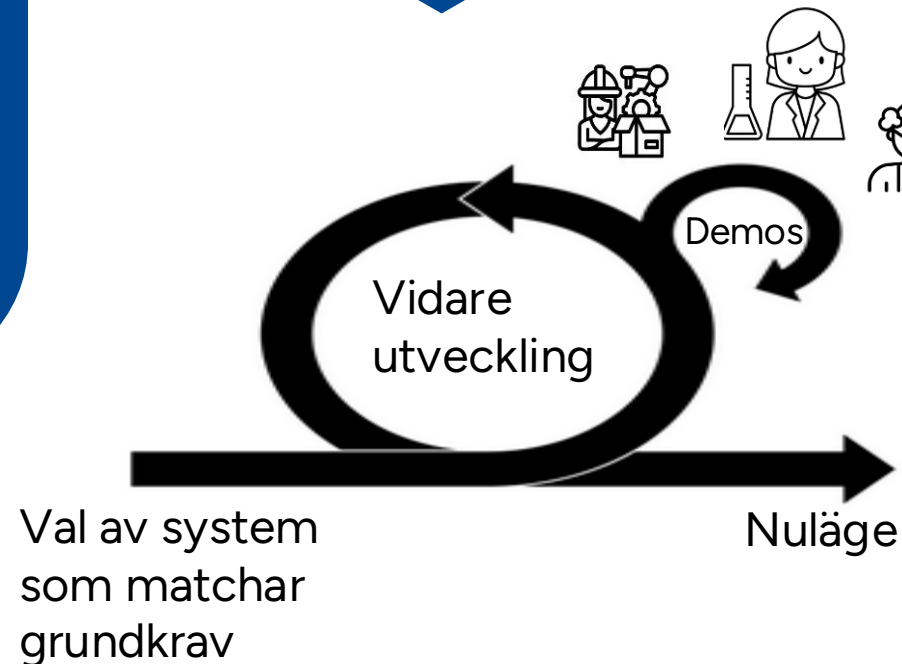
- Lagra data på KTH:s servrar
- Tilldela åtkomst till forskare från andra organisationer i mitt samverkansprojekt
- få en DOI för statisk version av källkod och data samt binders (kod/data/analys)

Etc, etc.....

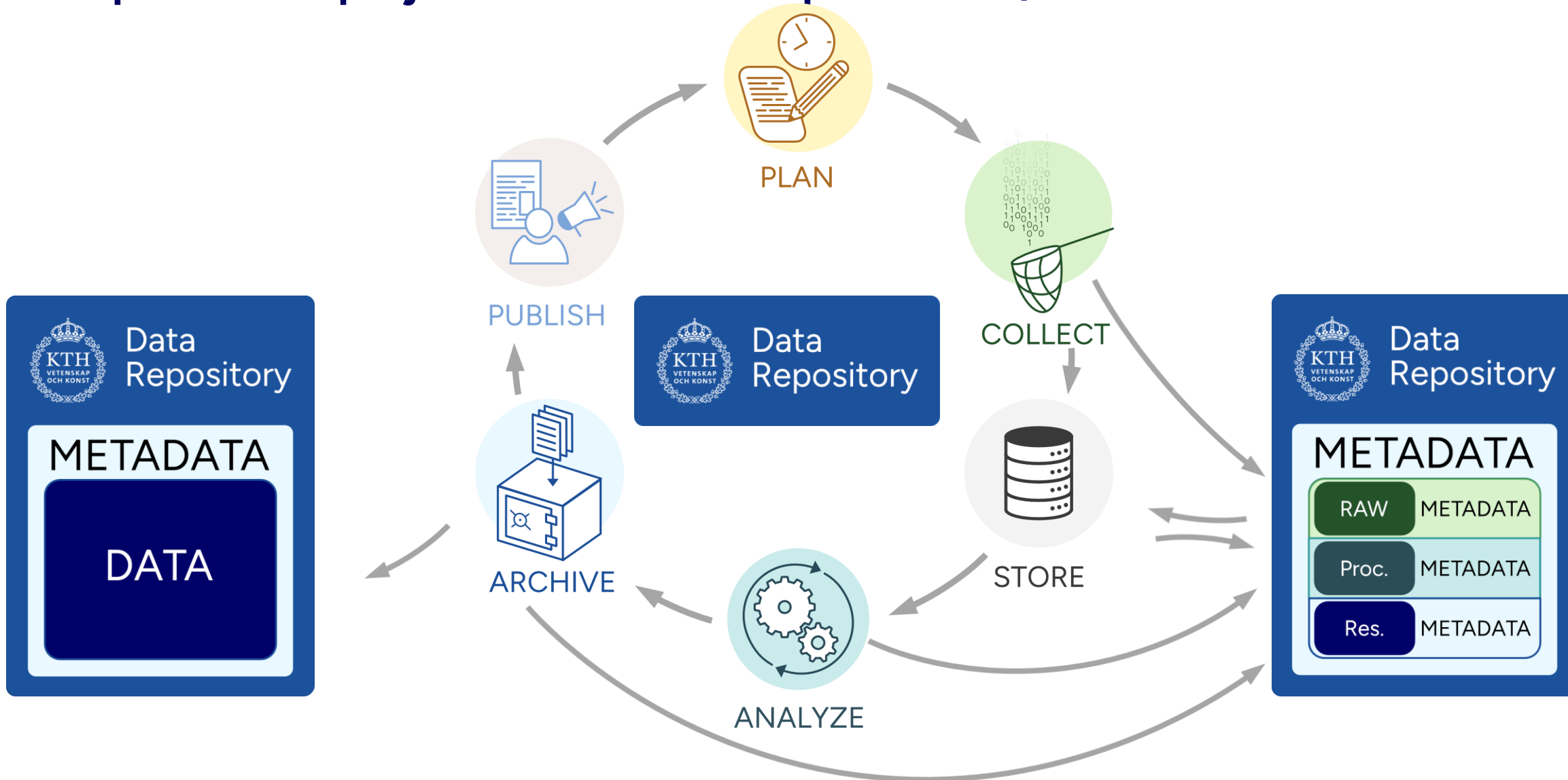
Status nuläge:

Vi har

- Ett system i produktion som matchar initiala krav samt tilläggskrav för forskare med "mellanstora behov" (K8S /S3)
- Genomfört initial RSA - förfining av icke-funktionella krav
- Dokumentation och utbildningsmaterial WIP
- Planerat lanseringsdatum 1:a januari



Ex. på arbetsflöden för forskare (utifrån input från forskare på demos i projekt för KTH data repository)



Invenio-projektet: förväntad nytta och effekter

Genom att införa systemstöd för att beskriva, tillgängliggöra och återanvända KTH:s forskningsdata och källkodbaserad metodik för KTH:s forskare så:

- Underlättas mer automatiserade informationsflöden och delning av data i samarbeten.
- Fås ett gränssnitt mot långsiktig lagring & backup av versionerade dataset som är användbart både för manuell och automatiserad åtkomst
- För publicerade datamängder möjliggörs länkning till publicerade resultat och källkod.
- Möjligheten att dokumentera både rådata och processad data med begränsad åtkomst stöttar forskare i att bedriva forskning på ett mer transparent och reproducerbart sätt

Systemstöd i kombination med insatser för att förtydliga ansvar för data och stöd kring domänspecifik god forskningspraxis för olika typer av data kan ge ytterligare effekter för KTH som organisation genom att:

- Forskningsdata nyttjas på ett mer effektivt sätt i forskningsprocessen
- En större del av KTH:s forskningsoutput synliggörs och bidrar till genomslag för KTH:s forskning
- På sikt få bättre kostnadskontroll över snabbt ökande datamängder som används och skapas i forskningen

