

Stockholm 2023-05-05

# Harmonisering av befintliga data hos datavärd: Hur kan dataproduktspecifikationer bidra till miljömålsuppföljningen

Finn Baumgartner





# Datavärdar - vision

- Datavärdarna svarar för leveranskontroll, lagring och presentation av data.
- Alla data som samlas in med hjälp av skattemedel är allmänt tillgängliga
- Data används och kommer till nytta.
- Alla data samlas in med gemensamma metoder.
- Alla data är tillgängliga via webben.
- Kvalitetsinformation finns för alla data.



# Målbild för datavärddskap?

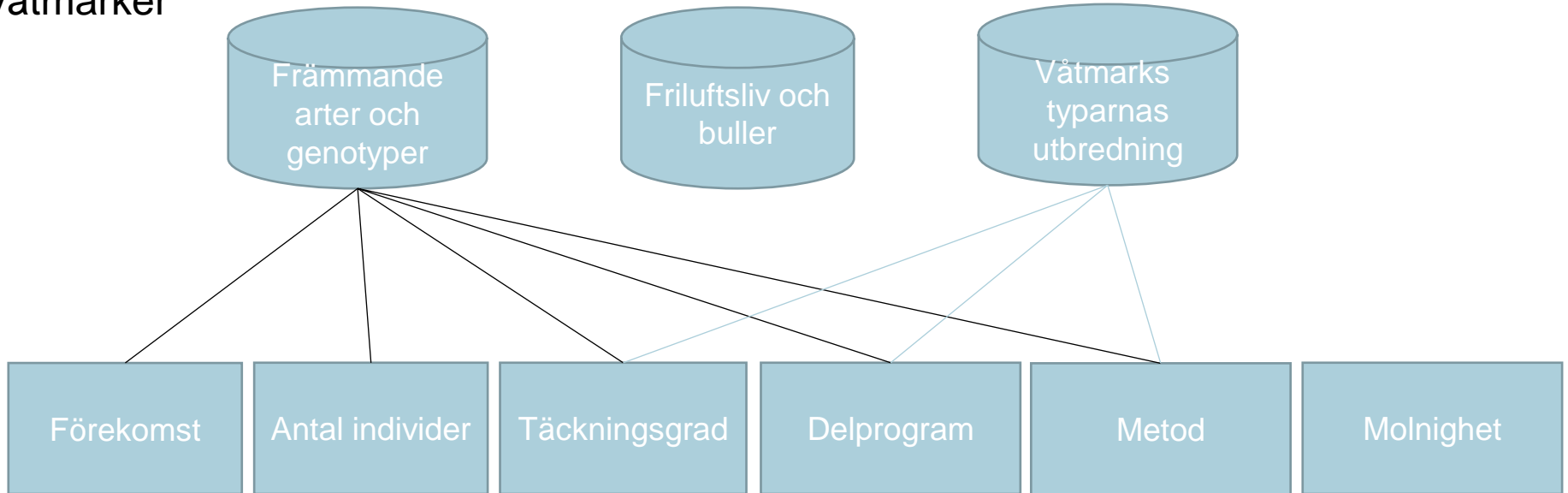
- På vilket sätt kan data göras lättbegripligt?
- Hur ska data kunna kombineras?
- Hur kan data bli lätt att använda?
- Vilka konsekvenser får det för bakomliggande system och informationsflöden?
- Vad behövs rent datamässigt för att kunna bidra till en uppföljning?

➤ *Frågor som behöver arbetas med i samverkan och med kompetenser från berörda intressenter*

# Vilka data behövs?

Myllrande  
våtmarker

?



# Dataproduktspecifikation (DPS)

$$\begin{aligned}\sin x &= 0.5 \\ \sin^{-1} \sin x &= \sin^{-1} 0.5 \\ x &= \sin^{-1} 0.5 \\ x &= 30^\circ\end{aligned}$$



$$\begin{aligned}\sin x &= 0.5 \\ x &= 30^\circ\end{aligned}$$

- Detaljerad beskrivning av en datamängd eller serier av datamängder tillsammans med ytterligare information som möjliggör att densamma kan skapas, levereras och användas av tredje part
- Kan produceras av olika parter för olika syften t.ex.:
- Styra datainsamlingsprocessen
- Skapa produkter från befintliga data

# Exempel process för DPS

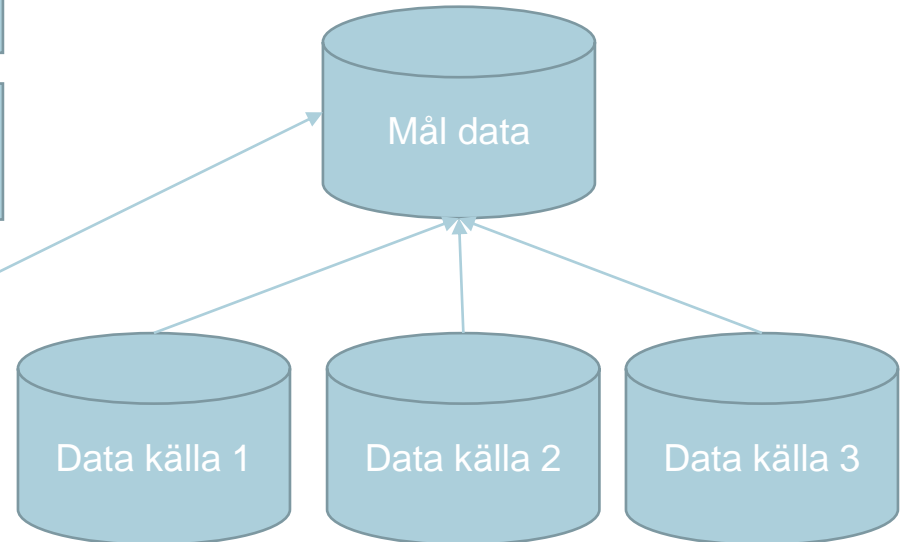
Verklighets baserad användningsfall

Analys av data källor

Identifiera behov av data harmonisering

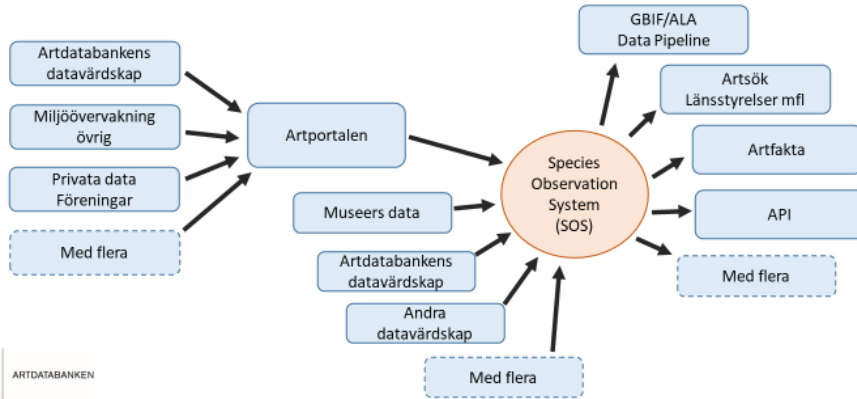
Skapa en gemensam data modell

Specifikation för mål data



Fichtinger et al. 2011. Data Harmonisation Put into Practice by the HUMBOLDT Project  
International Journal of Spatial Data Infrastructures Research, 2011, Vol.6, 234-260

# Data harmonisering



- Att kunna utbyta och/eller visa data tillsammans från olika källor måste mål data vara harmoniserad
- Bäst att det sker tidigt i dataflöden t.ex. ingår i beställning av datainsamling som styrs av en dataproduktspecifikation
- Inte alltid möjligt och inte fallet med befintliga data
- Oftast inte möjligt/praktiskt att harmonisera allt, kräver prioritering

# Gemensam språk

- För att kunna kombinera data från olika datakällor måste man ensa kring betydelsen av olika termer i ett särskilt sammanhang

Några förkortningar som kan ha olika betydelser beroende på sammanhanget:

- **SOS:**

Save our souls, Species observation service, Sensor observation service

- **SIS:**

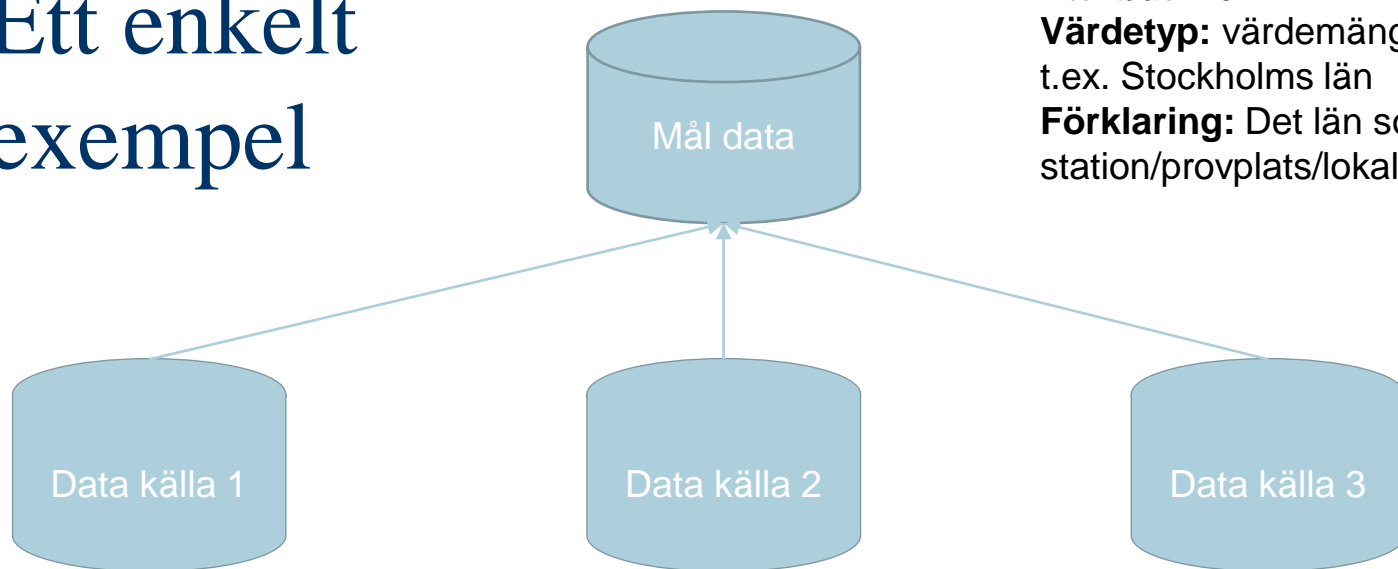
Svenska institutet för standarder, Seafloor Information Systems

- **FB:**

Facebook, Finn Baumgartner



# Ett enkelt exempel



**Attribut:** Län

**Värdetyp:** värdemängd, Län  
t.ex. Stockholms län

**Förklaring:** Det län som en station/provplats/lokal ligger inom

Data källa 1

**Attribut:** Län

**Värdetyp:** värdemängd,  
ISO 3166-2 kod

t.ex. SE-AB = Stockholms län

**Förklaring:** Det län  
som provplatsen ligger  
inom.

Data källa 2

**Attribut:** Lst

**Värdetyp:** värdemängd,  
länskod

t.ex. 01 = Stockholms län

**Förklaring:** Det län som  
stationen ligger inom.

Data källa 3

**Attribut:** Länsstyrelsen

**Värdetyp:** värdemängd,  
Län

t.ex. Stockholms län

**Förklaring:** Det län  
som en lokal ligger  
inom.

# Dataproduktspecifikation artobservationer

- Framtaget av NV tillsammans med datavärddar inom naturdata. Lunds universitet, SLU Adb, SLU skoglig resurshushållning
- Inkludera fält/attribut som är gemensamma för de tre datavärdarna
- Hantera inte alla fält/attribut som samlas in inom en viss datamängd
- Används som grund för att skapa en gemensam API specifikation
- Länk: [datavårdskap \(naturvardsverket.se\)](https://naturvardsverket.se/datavardskap)



2023-04-05

DATAPRODUKTSPECIFIKATION FÖR TILLGÄNGLIGGÖRANDE

## Naturdata – Systematiska artobservationer

Tabell 1: Versionsuppgifter

| Versionsuppgift           | Värde   |
|---------------------------|---|
| Version av specifikation: | 1.0   |
| Denna version:            | <a href="https://www.naturvardsverket.se/vagledning-och-stod/miljoovervakning/nationella-datavardskap/#E314836159">https://www.naturvardsverket.se/vagledning-och-stod/miljoovervakning/nationella-datavardskap/#E314836159</a> |

# Sammanfattning

- Data är en produkt som behöver beskrivas
- Viktigt att identifiera en fråga som data kan svara på
- Identifiera datakällor och specifika fält/parametrar som är nödvändig för att svara på frågan
- Identifiera om något saknas
- Identifiera harmoniseringsbehov
- Kom överens om gemensam språk

Tack!

NATUR  
VÅRDS  
VERKET

