

Datakällor, dataintegration och verktyg - översikt

Jerker Moström



Sveriges officiella
statistik

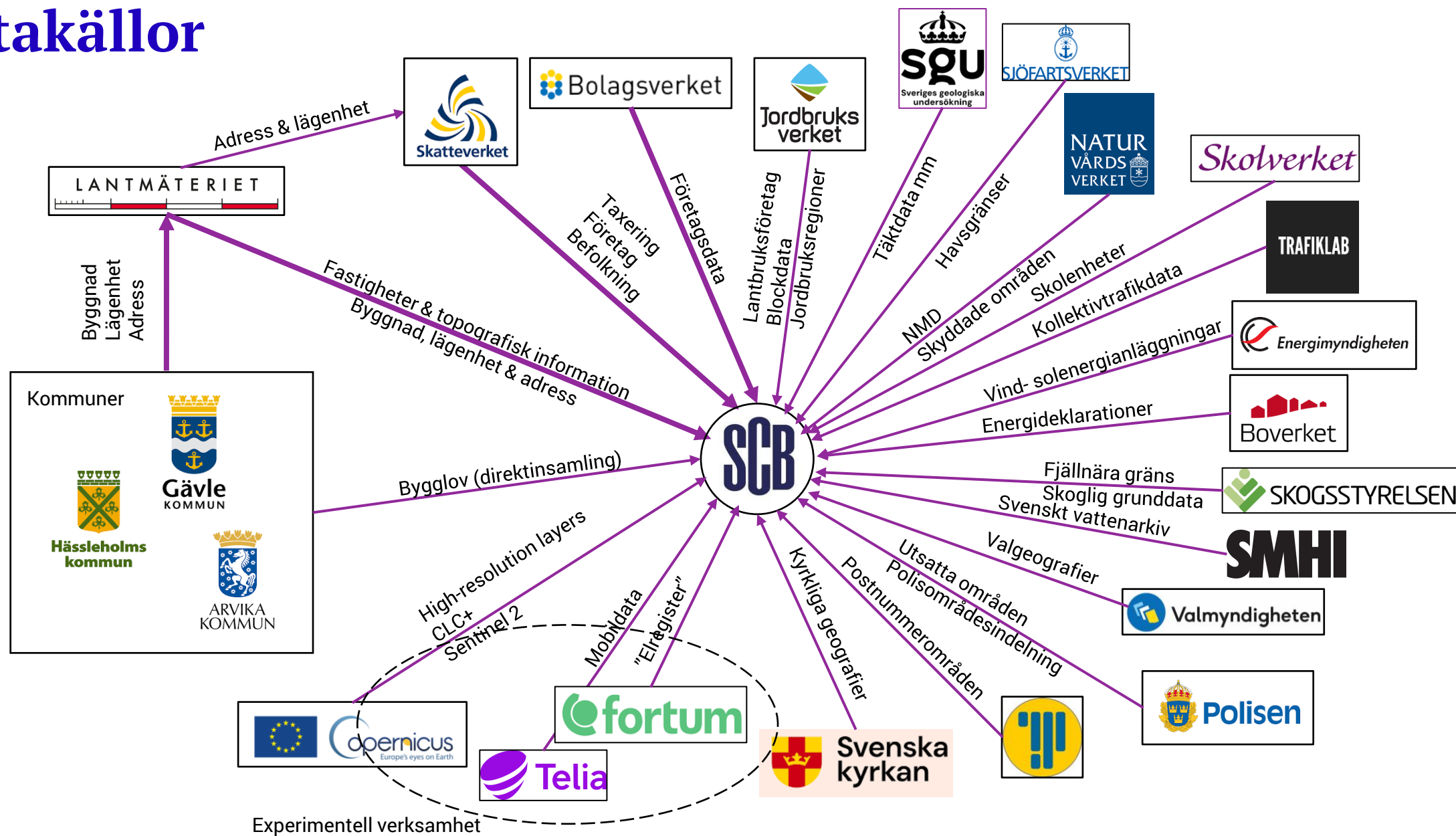
Pågående förändringsarbete

- Genomgripande förändring – lämna direktinsamling, använda mer befintlig data, tex administrativ eller annan data (gemensamt för hela statistikvärlden). Mer M2M.
- Minska bördan på uppgiftslämnare och sänka kostnaderna för insamling ("en uppgift en gång")
- Stort (och ökande) beroende av data från andra producenter (inte minst kommunerna!)
- Även data från privata aktörer (el, mobildata, kassaregister etc)
- SCB sitter inte på ett "eget förråd" av data

Geodata har olika funktioner

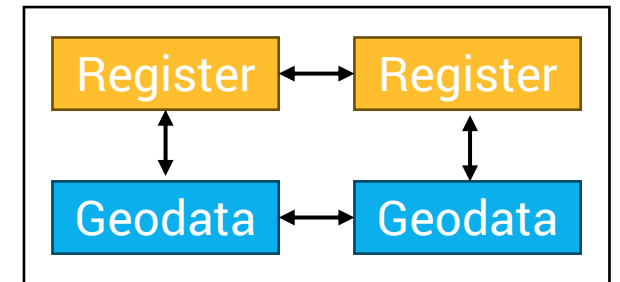
- **Stödjande**, del av den statistiska infrastrukturen (exempelvis adresser, administrativa och statistiska indelningar)
- **Statistiskt innehåll**, data som omvandlas till statistik (exempelvis markanvändning och marktäcke)
- Viss data har dubbla funktioner, exempelvis byggnadsinformation.

Datakällor

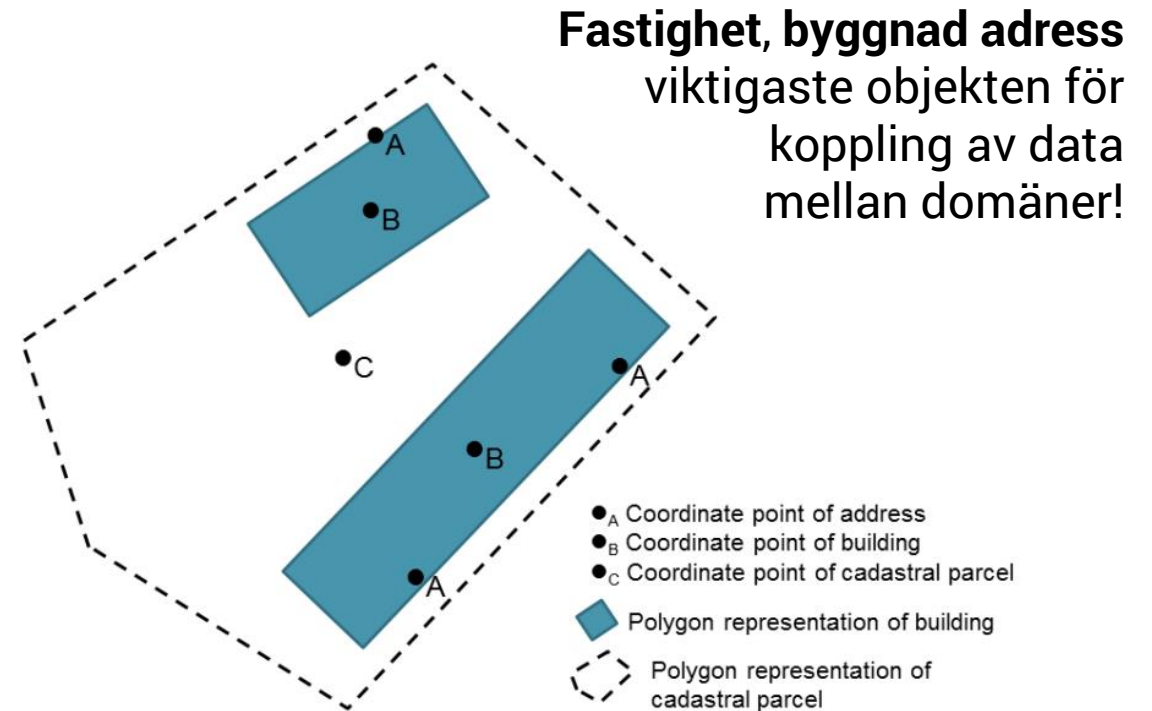
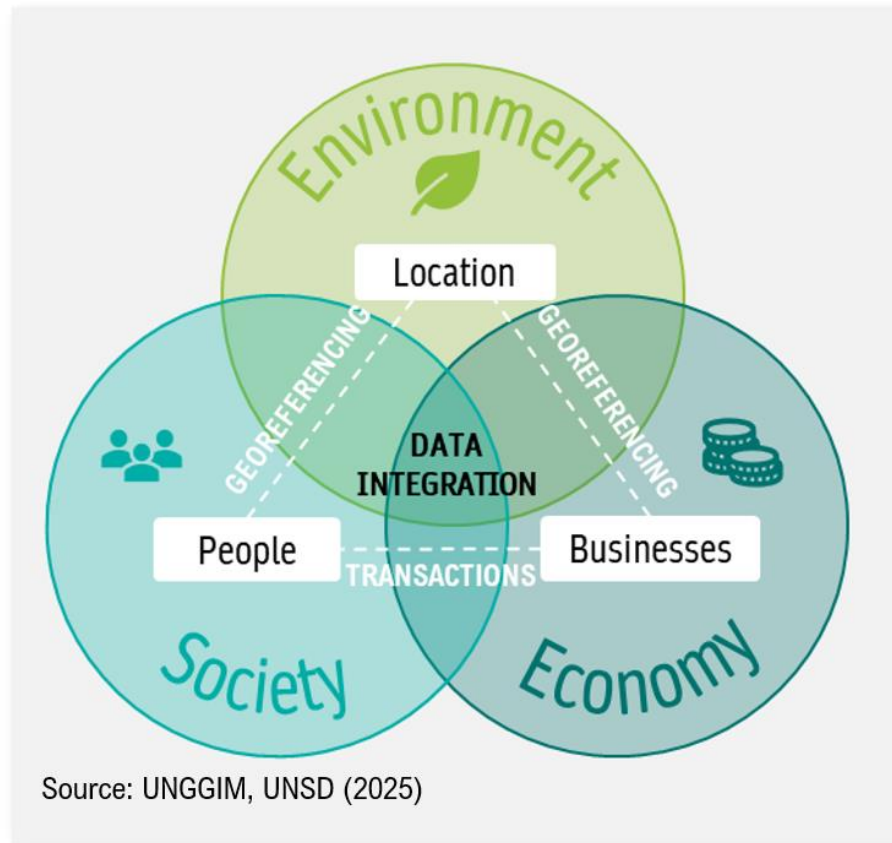


Statistiken är utgångspunkter för vår datahantering

- **Vad?** ("statistikvärde")
- **När?** (referenstidpunkt, tidsstämpling)
- **Var?** (geografi)
- Vi behöver frysa saker i tid och rum och spara/lagra väldigt mycket data (årsversioner)
- Arbetar under statistiksekretess
- Sammankoppling av data är centralt (matchning och geokodning)

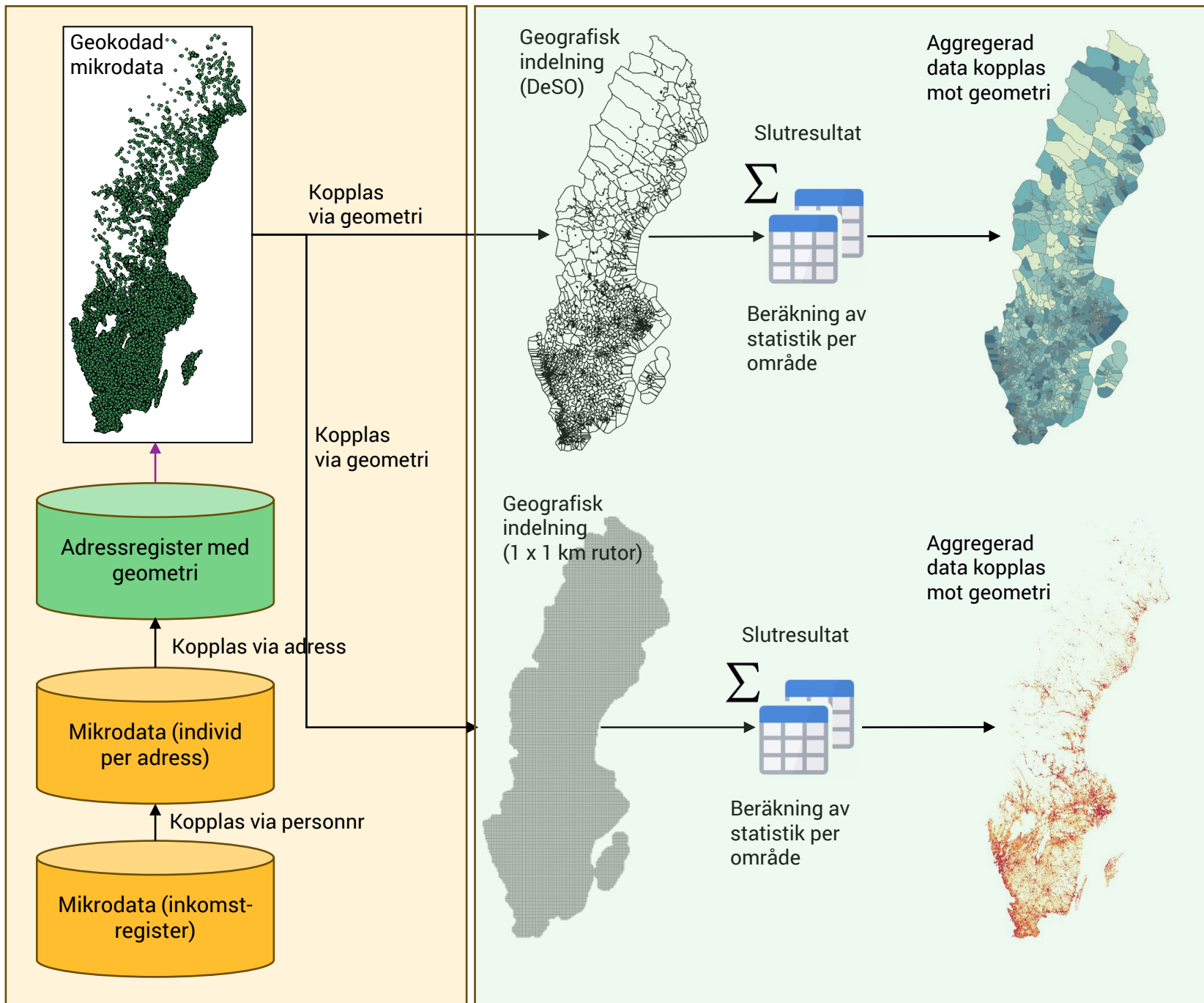


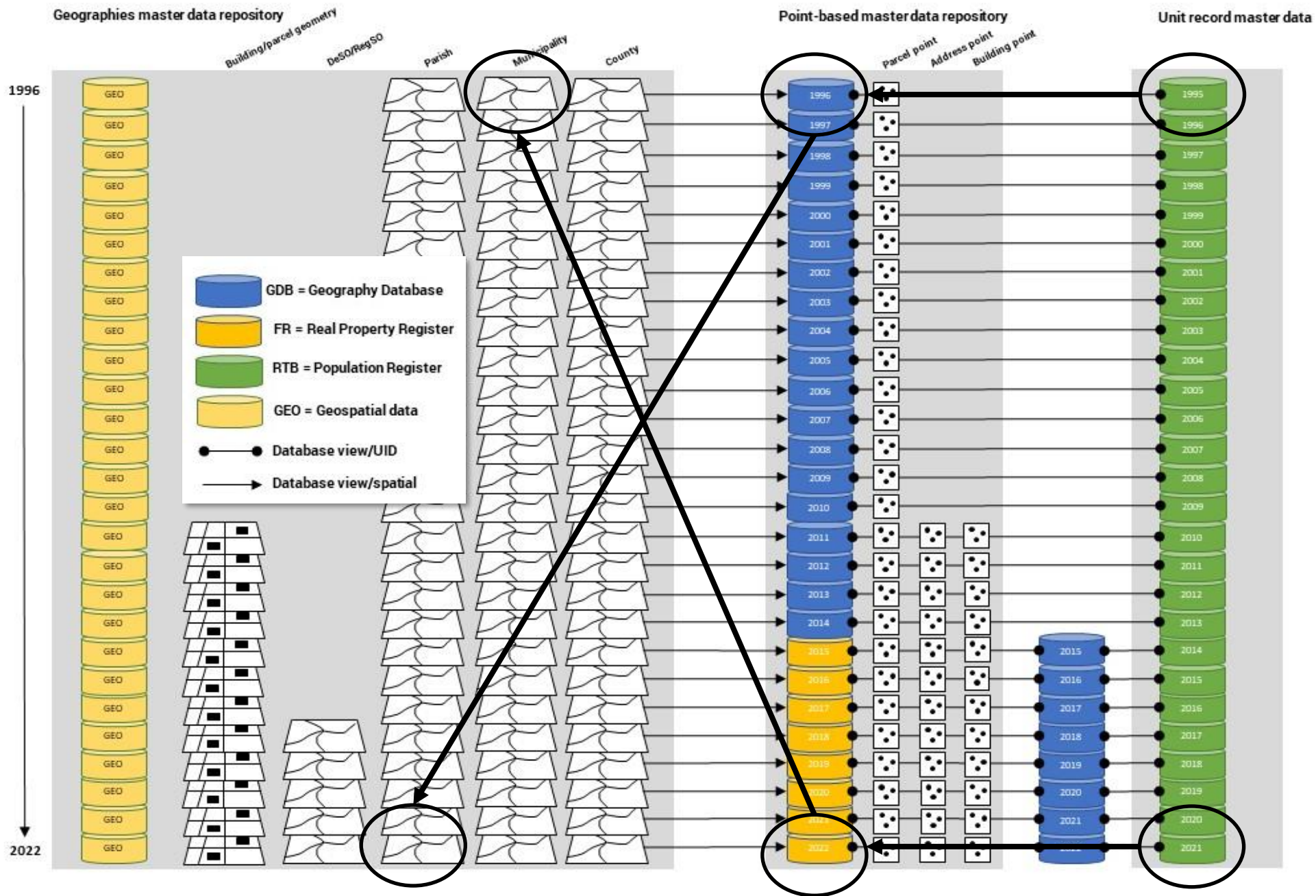
Dataintegration



Inom
SCB

Utanför
SCB





Verktyg (1)

FME

- Uppkoppling/hämtning av data via API, komplexa processer, bearbetningsflöden och geobearbetning

MS SQL server

- Spatial SQL för integration av geodata och registerdata som ej kräver geobearbetningar (typ "point-in-polygon")

Geoserver

- Publicering externa geodatatjänster (inklusive INSPIRE visning)

Verktyg (2)

MapInfo

- Produktionsstöd, kartproduktion

ArcGIS Pro

- Rasterbearbetning, editering, geometritvätt, layout/kartproduktion, publicering ArcGis Online

QGIS

- Breddanvändning (används även utanför "GIS-kretsen")

Översyn pågår kring mer ändamålsenlig infrastruktur för geodata

Utmaningar

- Datakvalitet!
- Brist på interoperabilitet (oftast semantisk)
- Geodatahantering i en värld av statistik
- Informationssäkerhetsklimatet hårdnar (begränsar tillgång till data)

Nationella geodataplattformen (NGP)

- Steg mot "dataekosystem-tanke"
- SCB håller i arbetet med temat Administrativa & statistiska indelningar
- Deltar även i arbetet kring Marktäcke & markanvändning
- Löser vissa interoperabilitetsproblem och ökar harmonisering

Administrativa och statistiska indelningar

= Påbörjat

