

Rapport

Beskrivning av datakatalog

ES2024-10





Innehåll

1.	Inledning.....	4
1.1	Målgrupp	4
1.2	Bakgrund	4
1.3	Syfte.....	4
1.4	Ambition.....	4
1.5	Användning.....	4
1.6	Grundläggande begrepp och termer	5
1.7	Begreppsmodell.....	6
2.	Datakatalog.....	7
3.	Datakatalogens innehåll.....	8
3.1	Gränssnitt.....	8
3.2	Gränssnittstyper	8
3.3	Funktioner.....	8
3.4	Behörighetsstyrning	9
3.5	Metadatahantering för datakatalog.....	10
3.5.1	Metadatahantering	10
3.5.2	Begreppsmodell metadatahantering	11
3.5.3	Metadatatyper.....	12
3.5.4	Metadainnehåll.....	14
3.5.5	Kvalitativ metadata	16
3.6	Informationssäkerhetsklassificering av metadata	17
3.7	Gemensamma metadatastrukturer.....	19
3.8	Specifikationer i datakatalogen.....	21
3.8.1	Specifikationer som dokument	21
3.8.2	Specifikationer som generella metadataprofiler.....	21
3.8.3	Specifikationer som enskilda metadataprofiler.....	22
4.	Referenser	23



1. Inledning

1.1 Målgrupp

Denna beskrivning riktar sig i första hand till arkitekter inom en organisation som avser att utveckla sin datahantering nyttjande en datakatalog för att effektivare kunna katalogisera samt dela data inom och utanför den egna organisationen.

1.2 Bakgrund

Hantering av data är fragmenterad inom organisationer i Sverige. Ökade krav på tillgång till data driver oss myndigheter mot en mer harmoniserad syn på hur data ska hanteras och även kunna utbytas mellan organisationerna. I ljuset av detta så har eSam tagit fram en referensarkitektur för *Hantering av data och utbyte av data mellan myndigheter* (se kapitel 4, referens 2) där datakatalog en av de centrala komponenterna.

1.3 Syfte

Syftet är att ge stöd och förståelse avseende för innehållet i en datakatalog för vidare implementering för att stödja katalogisering samt möjliggöra sökbarhet bland dataprodukter. Detta för att möjliggöra en effektiv och säker tillgång till data inom en organisation samt mellan organisationer inom Sverige, exempelvis mellan myndigheter. Denna beskrivning tar inte ställning till val av produkt, kommersiell eller ej.

Rapporten är skriven av Jan Aspenfjäll (Riksarkivet), Peter Hammar (Tullverket) och Dan Stenvall (Tullverket).

1.4 Ambition

Ambitionen med beskrivningen av datakatalog är att stödja referensarkitekturen Datahantering och utbyte av data mellan myndigheter (se kapitel 4, ref 2) samt skapa förståelse och ge guidning inför ett införande av en datakatalog.

1.5 Användning

Detta dokument är en av flera produkter inom eSam som berör området data. Referensarkitekturen är en rekommendation som är till för att användas av myndigheter inom eSam för framdrift av området Datakatalogen är tänkt att användas som en



komponent (se referensarkitektur Datahantering och utbyte av data mellan myndigheter kapitel 4, ref 2) och den gemensamma källan för att kunna publicera, administrera, hitta, söka och tillgängliggöra en organisations normaliserade och paketerade datatillgångar. Se också kapitel 4, referens 3.

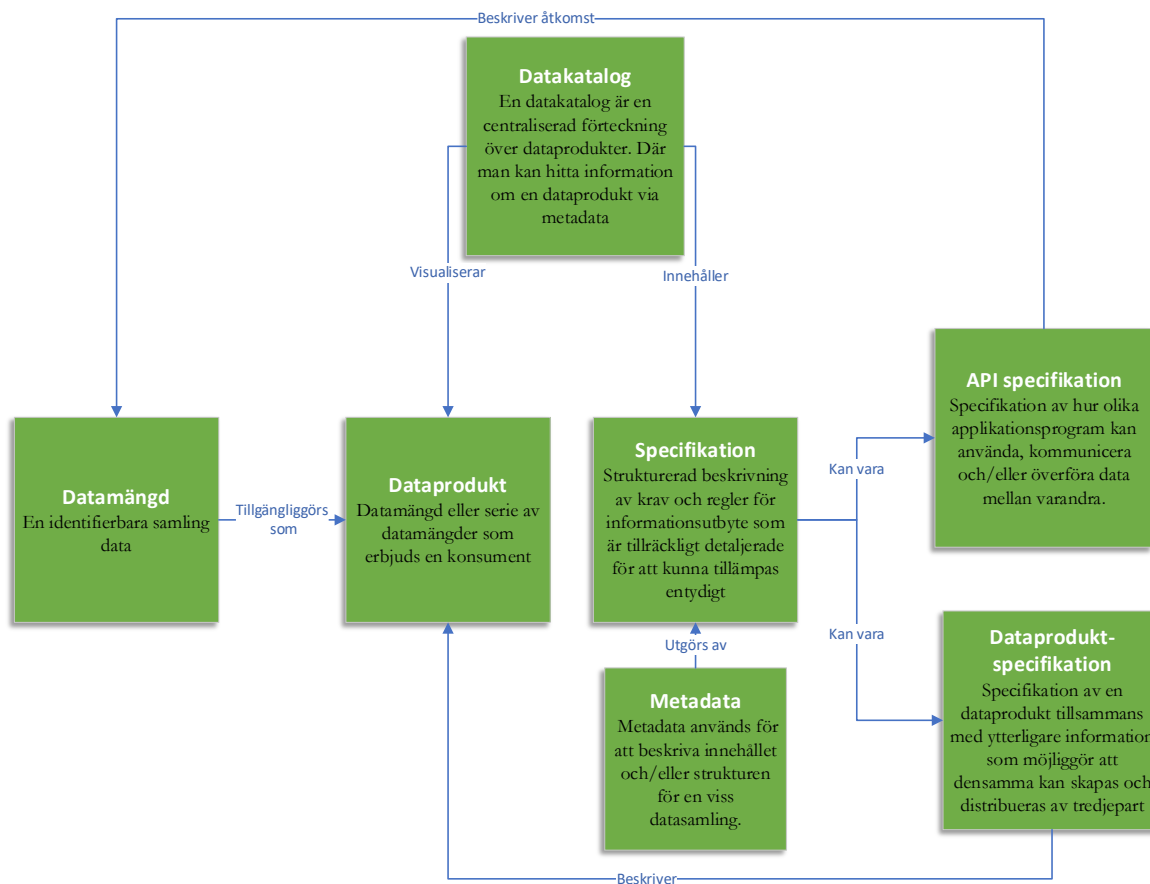
1.6 Grundläggande begrepp och termer

Nedan är ett uraval av viktiga begrepp använda i dokumentet. För en mer heltäckande lista se kapitel 4, referens 3.

- *API*: Application Programming Interface är specifikation av hur olika applikationsprogram kan använda, kommunicera och/eller överföra data mellan varandra
- *Data*: Data är råa, opolerade fakta och symboler som representerar händelser, objekt eller mätvärden
- *Datakatalog*: En datakatalog är en produkt som lagrar metadata om en datamängd
- *Datakälla*: Ursprunget för data, var datat har skapats
- *Datamängd*: Identifierbar samling data.
- *Dataprodukt*: Datamängd eller serie av datamängder som erbjuds en konsument
- *Dataproduktspecifikation*: Specifikation av en dataprodukt tillsammans med ytterligare information som möjliggör att densamma kan skapas och distribueras
- *Metadata*: Metadata används för att beskriva innehållet och/eller strukturen för en viss datasamling. Detta kan göras ur mer än ett perspektiv.
- *Referensarkitektur*: En modell eller ramverk som beskriver en rekommenderad struktur och organisering för ex system inom ett specifik område. Den fungerar som en vägledning och standardiseringsgrund.
- *Skörda data*: Data harvesting är en process som läser/hämtar datamängder och/eller deras metadata mellan producent och konsument
- *Tillgängliggöra*: Ge tillgång till data. Oavsett om det görs frivilligt eller på grund av en skyldighet i en författning



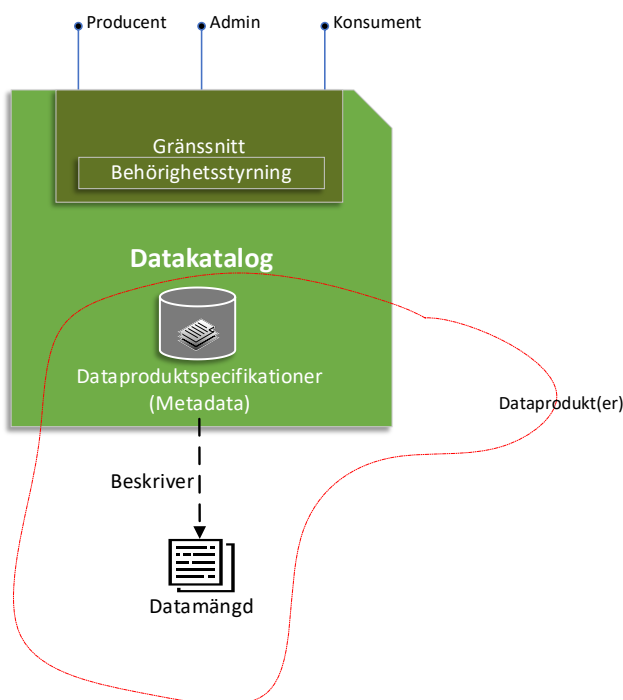
1.7 Begreppsmodell





2. Datakatalog

En datakatalog är en centraliserad förteckning över dataprodukter (ett kartotek) som finns tillgängliga i en organisation eller mellan organisationer. I en datakatalog kan man hitta information om en dataprodukts metadata så som information om dataproduktens namn, beskrivning, källa, tillgänglighet, datakvalitet, informations säkerhetsklassificering och användning, men också metadata om en mängd olika dataresurser, som databaser och API:er. Relationen mellan en datakatalog och en dataprodukt är att en datakatalog fungerar som en inventering av alla dataprodukter som finns tillgängliga. Datakatalogen innehåller dataproduktens specifikation "DPS" (eller referens till den) baserad på metadata. Datakatalogen behöver inte nödvändigtvis innehålla dataproduktens datamängd. Genom att ha en uppdaterad och korrekt datakatalog kan man lättare hitta och använda dataprodukter på ett effektivt sätt och man kan säkerställa att användningen av dataprodukterna sker på ett korrekt och säkert sätt.



En välskött datakatalog är en viktig resurs för att underlätta dataanvändning, säkerställa datakvalitet och efterleva regler och bestämmelser som gäller för datahantering och dataskydd. Det hjälper också organisationer att öka transparensen kring sina dataresurser och främja samarbete mellan olika avdelningar och användare som behöver tillgång till data. Eftersom datakataloger kommer att uppstå hos de flesta myndigheter så är det viktigt att man har interoperabilitet mellan dessa.



3. Datakatalogens innehåll

En datakatalog är en programvara (komponent) som måste kunna administreras och utvecklas efter organisationens behov. En datakatalog innehåller följande huvudområden

3.1 Gränssnitt

Datakatalogen har vanligtvis tre grundläggande gränssnitt:

- Gränssnitt för producent
- Gränssnitt för konsument
- Gränssnitt för Admin

3.2 Gränssnittstyper

Datakatalogen kan ha följande gränssnittstyper:

- Webb/GUI
- API

Webb/GUI byggs baserat på API:er. En rekommendation är att man identifierar och etablerar önskade API:er först. När dessa är på plats kan man sedan välja att exponera dessa direkt men även bygga web/GUI baserat på API:erna.

3.3 Funktioner

En datakatalog ska stödja ett antal funktioner för att kunna administrera, understödja och presentera dataprodukt, exempelvis:

- Funktioner för presentation av dataprodukt via Webb/GUI
- Funktioner för att söka data och dataprodukt
- Funktioner för att skapa sökfilter; användare kan använda filter och taggar för att förfinna sina sökresultat och identifiera dataprodukt som uppfyller specifika kriterier eller ändamål.
- Funktioner för att skapa, ta bort, uppdatera och publicera dataprodukt
- Funktioner för att importera och exportera dataprodukt, vilket inkluderar export till extern part



- Funktioner för hantering av metadata
- Funktioner för notifieringar om händelser i datakatalogen
- Funktioner för loggning och åtkomst till användningsdata (loggdata)
- Funktioner för åtkomstkontroll. Beroende på organisationens krav kan datakatalogen ha en integrerad åtkomstkontroll för att säkerställa att endast behöriga användare har tillgång till vissa dataprodukter
- Funktioner för hanteringsstöd av semantiska perspektiv på metadata samt import av metadatalistor
- Funktioner för datakatalogsstrukturer. Exempelvis olika kategorier eller ämnesområden baserat på de olika dataprodukterna eller affärsområdena de tillhör.

3.4 Behörighetsstyrning

Dataprodukten och dess delar, dataproduktspecifikationen samt datamängden kan behöva skyddas beroende på dess informationssäkerhetsklass. Man kan med fördel använda sig av datakatalogens metadata som en attributkälla för behörigheter och behörighetsprofiler. Syftet är att få en mer dynamisk behörighetsmodell som liknar ett filter och ger tillgång till de dataprodukter som matchas av värdena i dataproduktens attribut, samt vid utveckling metadatamodellen.

Skall datakatalogen användas för internt bruk så kan man i de flesta fall nyttja det interna behörighetssystemet. Kommer man sedan på att man vill dela dataprodukter externt så kan man överväga om man skall skapa en identitets och behörighets federation i syfte att minska administrationen av användare.

Nedan tabell ger exempel på funktion, roller och dess behov av behörigheter.

Funktion	Producent	Konsument	Admin
Skapa dataprodukt	x		
Uppdatera dataprodukt	x		
Aktivera/deaktivera dataprodukt	x		
Ta bort dataprodukt			x
Söka bland dataprodukter	x	x	x
Se statistik för dataprodukten	x	x	x
Skapa behörighetsprofil för dataprodukt	x		



3.5 Metadathantering för datakatalog

Metadata är data som beskriver andra data. Metadata ger information om en viss datamängd som kan hjälpa till att hitta, förstå och tolka dessa data. Metadata kan innehålla en mängd olika uppgifter, t.ex. vem som har skrivit ett dokument, vilket datum det skapades, filformat, storlek och var filen finns. I detta dokument har vi fokuserat på metadata för datakatalogen och behovet som varje organisation har för att kunna nyttja en datakatalog för att utbyta information internt eller med andra organisationer på bästa sätt.

Den sammantagna och paketerade mängden metadata som beskriver en datamängd kallar vi här för DPS (dataproduktspecifikation). DPS:en skall hållas separerad från den datamängd som den beskriver beroende på att den ofta är mer statisk än den datamängd som den representerar.

För att DPS:n ska ge så hög nytta som möjligt så ska den enkelt kunna konsumeras både visuellt och maskinellt. För att metadata ska vara användbart är det viktigt att de enkelt kan läsas både av producenter, användare och av IT-system.

Datakatalogen är en central databas som innehåller DPS:er eller referenser till dessa. Det är en organiserad förteckning av dataprodukter som ger användarna en överblick över vilka dataprodukter som finns tillgängliga. DPS:en är grundfundamentet för en datakatalog och metadata är grundfundamentet för DPS:en. I det här dokumentet kommer fokus ligga på hur vi med metadata organiserar och bygger upp en DPS.

Metadata behövs på olika detaljeringsnivåer och det är därför viktigt att så tidigt som möjligt sätta principerna för metadata.

3.5.1 Metadathantering

För att kunna förstå och utbyta data mellan olika nivåer så krävs standardisering och den nationella metadathanteringen måste utgå ifrån och harmonisera med EU-nivån. Detta är den första nivån av metadathantering. Nästa nivå är den nationella som ska vara gemensam för hela den offentliga sektorn, den 3:e nivån är där Statliga myndigheter, kommunala myndigheter och regionala myndigheter har sina specifika metadata. Den 4:e nivån är den som har kallats sektorsspecifik där de myndigheter som agerar i samma område kan påverka de specifika metadata som finns i den specifika sektorn, i detta uppdrag använder vi oss av termen dataområden som mer beskriver de



informationsområden som olika myndigheter (statliga, regionala och kommunala) delar för att utföra sina respektive uppdrag, se bild nedan.

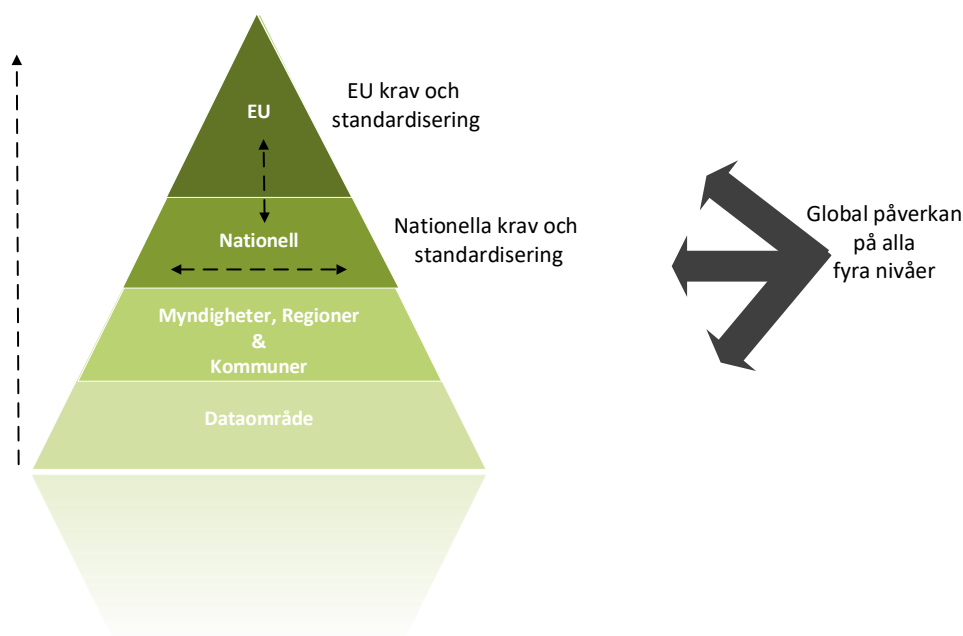


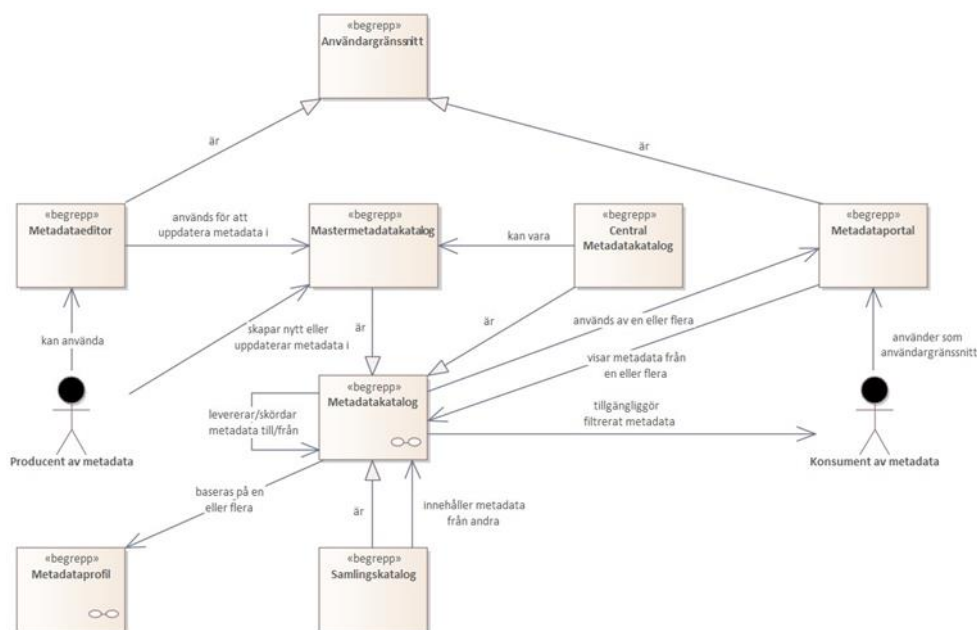
Bild över metadatahanterings olika nivåer.

3.5.2 Begreppsmodell metadatahantering

Det finns ett framtaget dokument som är en begreppsmodell för metadatahantering (se kapitel 4, referens 10) Denna begreppsmodell kan användas som ett referensramverk för att beskriva metadatamodellen för DPS:er som lagras i en datakatalog.

Varje organisation som ska ta fram en datakatalog går igenom de delar som finns med i begreppsmodellen och gör en bedömning av vilka delar det finns behov av och vad för funktion som måste finnas i respektive ingående del av begreppsmodellen.

Begreppsmodellen kan även användas för att ta fram modellbeskrivningar för attribut på metadatabegrepp (se kapitel 3.5.4).



Begreppsmodell metadata.

3.5.3 Metadatatyper

Metadata finns i olika typer och det finns en del olika sätt att beskriva dessa typer. Det vanligaste är att det finns tre huvudtyper av metadata: administrativ, beskrivande och strukturell. Vilket är valt i detta fall utifrån att den passar bra att beskriva de metadatatyperna som det finns behov av i en datakatalog.

Metadata måste delas upp i olika kategorier för att kunna stödja användaren på rätt sätt, t.ex. en utvecklare har behov av teknisk metadata för att kunna implementera metadata såsom datatyp, ev. längd på datatypen osv... medans en slutanvändare har behov av beskrivande metadata för att kunna fylla i en uppgift på rätt sätt (se kapitel 3.5.4).

3.5.3.1 Strukturella metadata

Strukturella metadata är den metadata som t.ex. behandlar navigation och presentation av elektroniska resurser. Strukturell metadata används t.ex. för att beskriva hur en databas ser ut, beskriva statistisk data samt geodata. Metadata beskriver hur behållare med data är strukturerad och anger hur sammansatta objekt sätts ihop, till exempel hur sidor ordnas för att bilda kapitel. Den beskriver typer, versioner, relationer och andra egenskaper hos digitalt material.



3.5.3.2 Administrativa metadata

Administrativa metadata är den metadata som används för att ange saker som teknisk data vid skapandet av resursen, information om ägare och rättigheter, hur man kommer åt och använder resursen samt hur den har och ska bevaras.

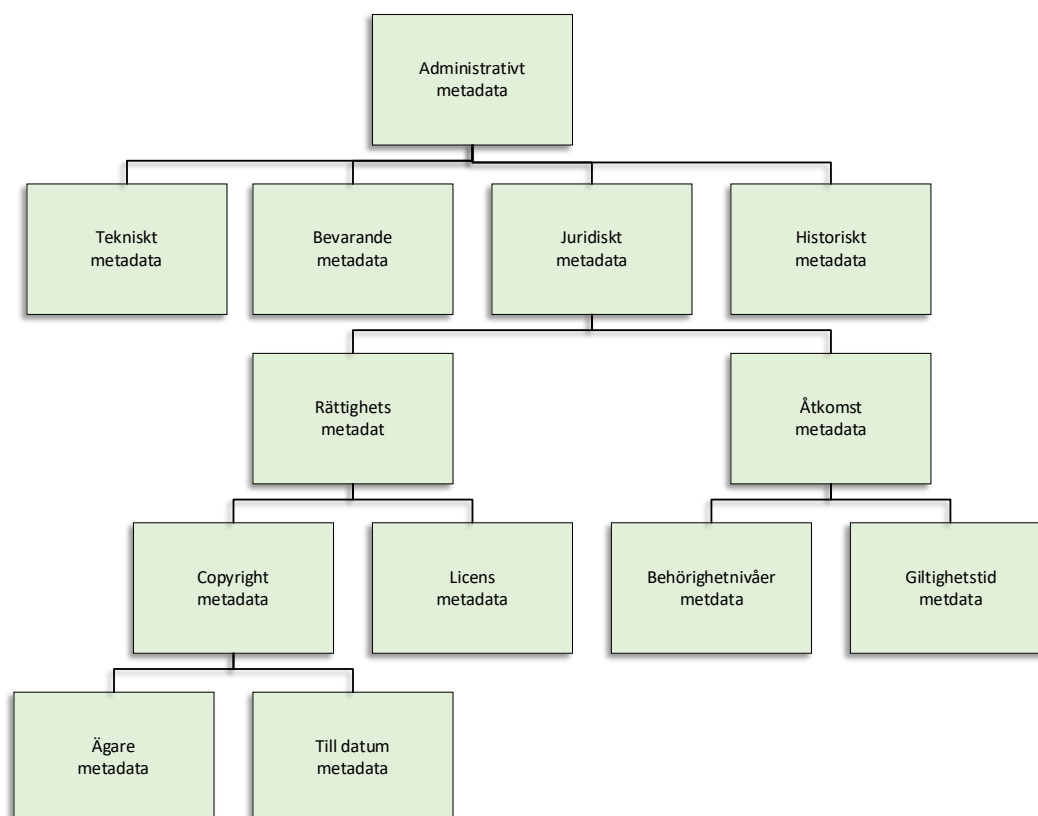
Administrativa metadata kan delas upp i underkategorier enligt följande:

- Teknisk metadata
- Bevarandemetadata
- Rättighetsmetadata
- Historisk metadata

Det kan finnas fler underkategorier beroende på vad organisationen har för behov.

3.5.3.3 Beskrivande metadata

Beskrivande metadata är den metadata som används för att beskriva och identifiera resurser genom t.ex. unika identifierare som namn och URI, fysiska attribut t.ex. typ av resurs, storlek och vikt och icke-fysiska attribut t.ex. titel, utgivare, ämne, nyckelord och beskrivning av resursen.



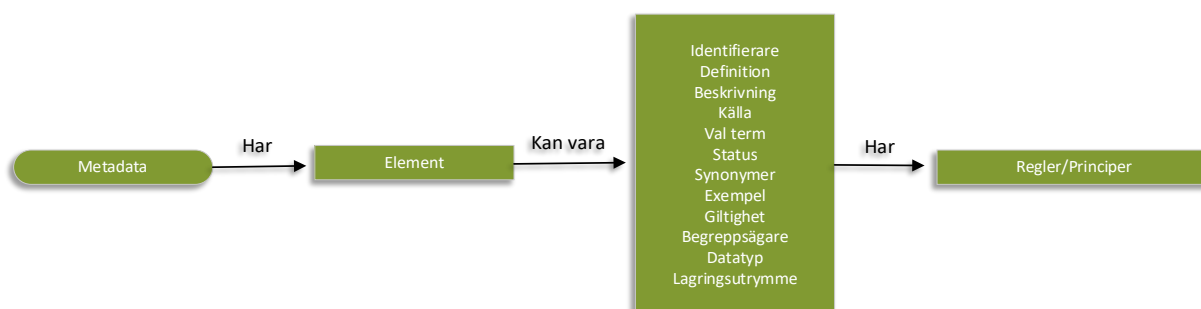
Exempel på hur det går att se strukturen på sin egna metadata.



För att förstå och få kontroll på organisationens metadata så går det att utifrån huvud metadata typerna att se på vilka undergrupper varje metadata typ som organisationen har behov av och sedan bryta ner strukturen tills man är på den exakta metadata man har behov av (se exempel i bilden ovan). När detta är klart så kan metadata informationssäkerhetsklassificeras på den nivå som behovet av klassificering finns på (se kapitel 3.6).

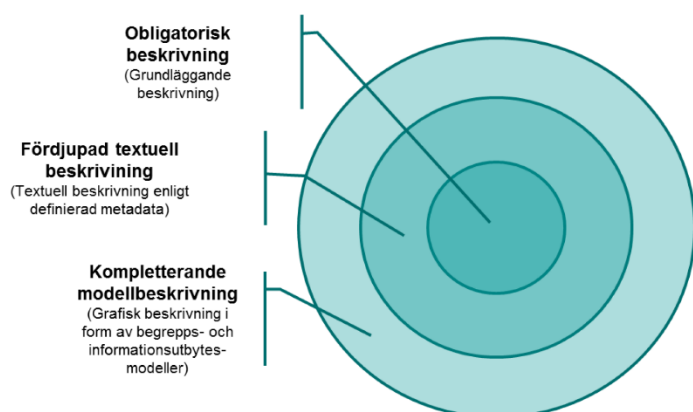
3.5.4 Metadata innehåll

Metadata ska ha en viss mängd information som gör den specifika metadata användbar både vad gäller sökbarhet, tillgängliggörande och för att möjliggöra teknisk implementering. Det bör beslutas vilka attribut som i de olika typerna av metadata är grundläggande textuell beskrivning (obligatoriska), fördjupad textuell beskrivning (vid behov) och kompletterande modellbeskrivning (frivillig), (se kapitel 4, referens 9 kap 4.1) detta dokument har några år på nacken men är fortfarande aktuellt.



Exempel på attribut på metadata.

Det finns några olika delar som beskrivs i processen och dessa är obligatorisk beskrivning, fördjupad textuell beskrivning och komplementär modellbeskrivning (se bild nedan). Dessa används för att beskriva de attribut som ett metadata begrepp har och hur metadata objektet kan användas utifrån olika användare.



Beskrivning av processens ingående delar.

De attribut som det finns behov av för metadata beskrivs i den obligatoriska beskrivningen och i den fördjupad textuell beskrivning. Attributen kan se olika ut även för en producent kan det vara olika beskrivningar beroende på om det är till en programutvecklare eller en verksamhetsutvecklare.

Den obligatoriska beskrivningen är på en övergripande nivå som ger en snabb förståelse för vad det är för metadata och hur den ska användas. (se bild nedan). Där finns i exemplet identifierare, namn, beskrivning och begreppsägare som attribut, dessa attribut har sedan en beskrivning och/eller regler knutna till sig i kolumnen som heter tänkt innehåll. Om det är en metadata som har behov av att beskrivas mer så finns det utrymme att göra det, bilden nedan är endast ett exempel på attribut som kan finnas med. Alla organisationer måste beskriva utifrån de behov som finns och hur väl mottagaren är med metadatabegreppet.

Egenskaper	Tänkt innehåll
Identifierare	Unik identifierare för begreppet (ska om möjligt vara en URI (Uniform Resource Identifier)).
Namn	Begreppets namn i klartext.
Beskrivning	Förtydliganden och förklaringar av ett begrepp, som inte är definition eller exempel.
Begreppsägare	Aktör som ansvarar för begreppet.

Exempel på attribut för den obligatoriska beskrivningen.

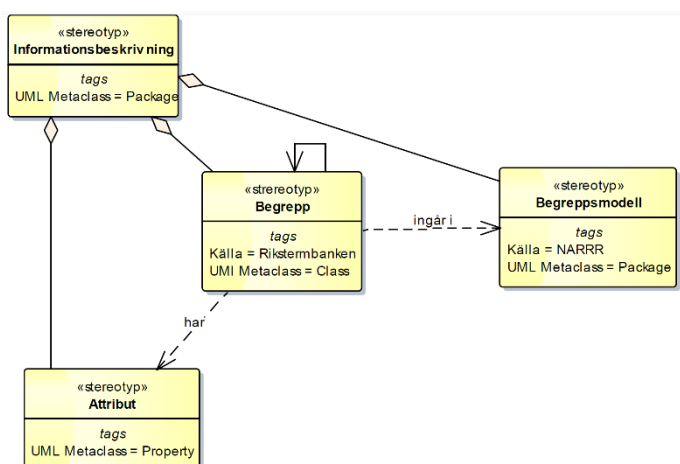
Den fördjupade textuella beskrivningen är den som fördjupar beskrivningen av metadata och som ska beskriva mer komplexa attribut för metadata. Här kan de mer komplexa attributen beskrivas, för exempelvis programutvecklare.



Egenskaper	Tänkt innehåll
Definition	Språkligt utformad beskrivning av ett begrepp som ska avgränsa detta mot andra begrepp.
Källa för definition	Anger ursprung till definitionen, exempelvis lagrum (lag, förordning, paragraf, stycke), organisation, Rikstermbanken.
Rekommenderad term för begreppet	Det ord som används för att tala eller skriva om begreppet.
Status	Vilken status begreppet har. Kan anta något av värdena: under utveckling, godkänd, inaktuell.
Giltighet	Från och till vilket datum begreppsbeskrivningen är giltig. Giltighet fylls i olika beroende på status: ”Under utveckling”: behöver inte giltighet fyllas i ”Godkänd”: fyll i giltig från ”Inaktuell”: fyll i giltig från och till
Om det finns synonymer till termen	Samtliga ord som har samma betydelse som vald term.
Exempel	Om det finns exempel som ökar förståelsen av begreppet.
Begreppsägare	Aktör som ansvarar för begreppet.
Länk till källa eller begreppsmodell	Hyperlänk.

Exempel på attribut på den fördjupade textuella beskrivningen.

Den komplimenterade modellbeskrivningen är den del för just detta metadatabegrepp som tagits fram i begreppsmodellen (se kapitel 3.5.2). Denna modellbeskrivning kan användas till att beskriva sammanhang och med hjälp av attributen göra det möjligt att särskilja det från andra likartade metadatabegrepp.



Exempel på hur en modellbeskrivning kan se ut för ett metadatabegrepp.

3.5.5 Kvalitativ metadata

Det finns i dagsläget ingen nationell modell som används gemensamt bland myndigheter för att bedöma en enskild metadataas kvalitet, detta är dock viktigt för att skapa en hög kvalitet på metadata och för att skapa ett effektivt utbyte av information via datakataloger. FAIR-principerna (se kapitel 4, referens 7) publicerades första gången 2016. Sedan dess har bland andra Europeiska unionen (EU) samt ett flertal forskningsfinansiärer, universitet och forskningsinfrastrukturer ställt sig bakom dem. FAIR principerna kan vara ett sätt att skapa en grund för metadatakvaliteten om den



används nationellt skulle det skapa en kvalitetsnivå som alla myndigheter (statliga, regionala och kommunala) kan verka runt för att utbyta data på ett effektivare sätt. Eftersom EU står bakom konceptet så skapar det en förståelse för kvalitén även vid utbyte av data inom EU.

FAIR principerna bygger på fyra grundstenar vilket är att metadata ska vara:

- Sökbara
- Tillgängliga
- Kompatibla
- Återanvändbara

Under dessa fyra grundstenar finns olika kriterier för att bedöma metadatakvalitén. Vi kan med hjälp av FAIR principerna nå en likvärdig kvalitet av metadata för myndigheter.

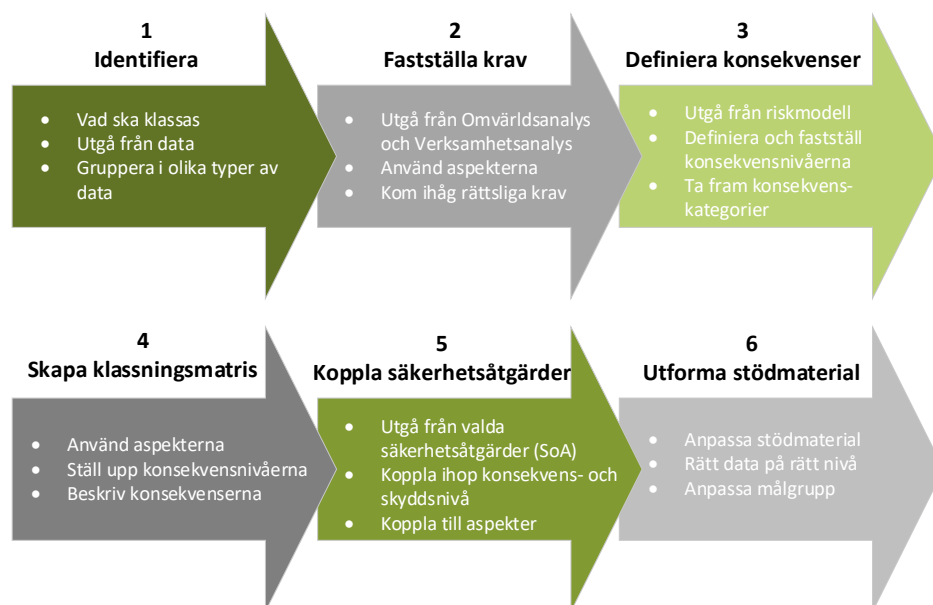
För att FAIR principerna ska vara effektivt att nyttja finns det applikationer för att testa dessa principer på metadatabegrepp och där se kvaliteten på metadatabegreppet utifrån de fyra grundstenarna.

3.6 Informationssäkerhetsklassificering av metadata

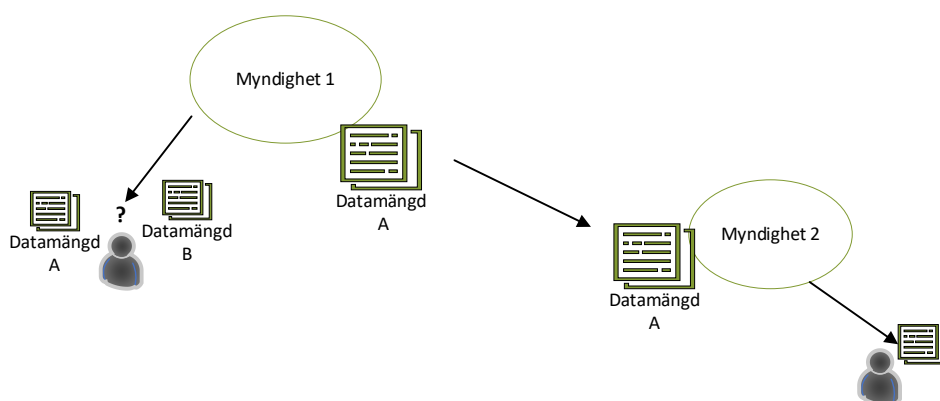
I dagsläget så informationssäkerhetsklassificeras data, men inte metadata i samma utsträckning. Metadatabegrepp som kan ”förstås” och göra så att det går att förstå vad data handlar om kan behöva informationssäkerhetsklassificeras. Även aggregerade metadata kan bli förståelig på grund av den samlade mängden metadatabegrepp och på så sätt kan data/information bli förståelig med endast metadatabegrepp. När metadata informationssäkerhetsklassificeras så bör alla myndigheter ha samma informationssäkerhetsklassificeringssystem så att klassificeringen blir förståelig. I dagsläget så sker informationsklassificering olika hos myndigheter på grund av olika bestämmelser för dessa. Myndigheten för samhällsskydd och beredskap (MSB) har en metod för informationssäkerhet. Denna metod borde testas för användning när det gäller metadatabegrepp. MSB har tagit fram en klassificeringsmodell som är en del av metodstödet och som kan ge en god vägledning för att klassificera data som kan fungera även för metadatabegrepp. Arbetet med att informationssäkerhetsklassificera en organisations metadata (men även datatillgångar) kan vara tidskrävande och kräver kunskap. Här finns stora möjligheter att nyttja AI-baserad teknik (kapitel **Fel! Hittar inte referenskälla.**, referens 5) eftersom tekniken är gjord för att kunna bearbeta



datamängder och med stöd av algoritmer och kunskapsmodeller så kan den ge ett effektivt stöd för informationssäkerhetsklassificering.



Informationssäkerhetsklassificeringsmodell inspirerad av MSB.

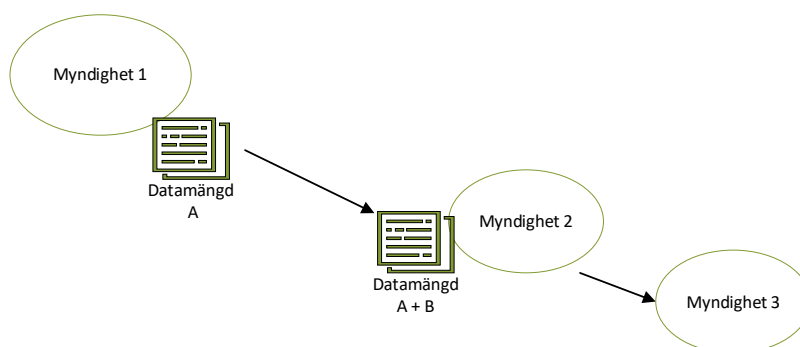


Exempel på säkerhetsprövning av data och metadata

Information med metadata som skickas mellan myndigheter och som blir inkommen hos en ny myndighet som gör en sekretess-/säkerhetsprövning enligt sin myndighetspolicy. Det som då sekretess prövas och säkerhetsprövas är själva data, men det finns ett behov av att även sekretesspröva och säkerhetspröva metadata eftersom viss metadata



säger mycket om datat. Hos myndighet 1 kan sekretess-/säkerhetsprövningen göra så att mindre data släpps ut än hos myndighet 2 där man har en annan policy vid sekretess-/säkerhetsprövningen. Själva metadata sekretess-/säkerhetsprövas inte alls, vilket skapar risk att det går att förstå vad data handlar om. Data som förändras, det vill säga det läggs till data eller det tas bort data, och med det även metadata. Denna typ av händelse måste hanteras samt skapas rutiner för.

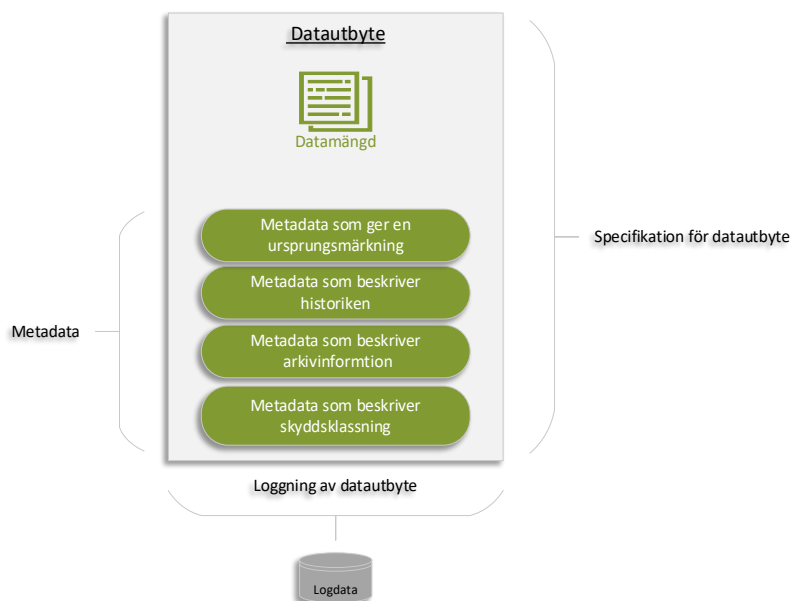


Exempel på när information förändras av myndighet 2

Om myndighet 1 skickar data till myndighet 2 och lägger till data och metadata hur förändrar de informations säkerhetsklassificeringsgraden av den nya datamängden? När den nya datamängden hamnar hos myndighet 3 så kanske, för den myndigheten, har datamängden ingen hög informations säkerhetsklassificeringsgrad och då sänkes informations säkerhetsklassificeringsgraden utifrån datamängdens tidigare informations säkerhetsklassificeringsgrad. Om det finns ett gemensamt sätt att klassificera metadata på samt principer att en informations säkerhetsklassificering inte får sänkas utan tillstånd av den myndighet som från början skapade data. Det är sådana principer som behövs för att skydda data så att det inte är lättare att få ut klassificerad information från vissa myndigheter och svårare från andra när det utbytes information mellan organisationer (se kapitel 4, referens 8). De är juridiskt möjligt i dagsläget att sänka skyddsklassningen.

3.7 Gemensamma metadatastrukturer

Det finns ett behov av att ha gemensamma metadatastrukturer vid utbyte av data. Detta för att se historiken, proveniens, arkivinformation och en gemensam informations säkerhetsklassificering. Det vill säga att det är ett sätt att säkra information och ger ett stöd till att förstå vilken metadata som det finns behov av vid olika tillfällen.



Exempel på vilken typ av metadata som det finns behov av för att säkerställa data

För att säkerställa ett datautbyte så finns det behov av loggning både för avsändare som för mottagare, dels se att det kommit fram till rätt mottagare och att avsändaren vet att det kommit fram till rätt mottagare. Detta behövs även för rättssäkerheten gällande att data inte kommer i orätta händer. Det är därför viktigt att ta fram vad behöver avsändare respektive mottagare för metadata vid loggning.

Avsändaren behöver kunna påvisa att rätt mottagare fått data och när den mottagits och hanterats av mottagaren. Mottagaren måste veta att data kommer från den avsändaren som de vill ha data ifrån. Detta är bara exempel på information som kan behövas och varje organisation har behov av för att klargöra vilken typ av metadata som den har behov av vid datautbyten med externa organisationer eller medborgare.

Det finns behov av att veta vilken historik data man har utifrån om detta är originaldata eller om den har förändrats på något sätt och då även få provenienser och ursprungsmärkning med för data. Finns det data som ska gallras eller bevaras och under vilka lagrum som det ska gallras eller bevaras så kan detta vara viktigt för mottagande organisation eller om detta ska överlämnas till arkivet så måste denna data finnas med.

När det gäller informationssäkerhetsklassificering så är det viktigt att denna information finns med och att det är relevant metadata så att mottagande organisation förstår informationssäkerhetsklassificeringen. Det måste finnas metadata som beskriver vad informationssäkerhetsklassificeringen innebär.

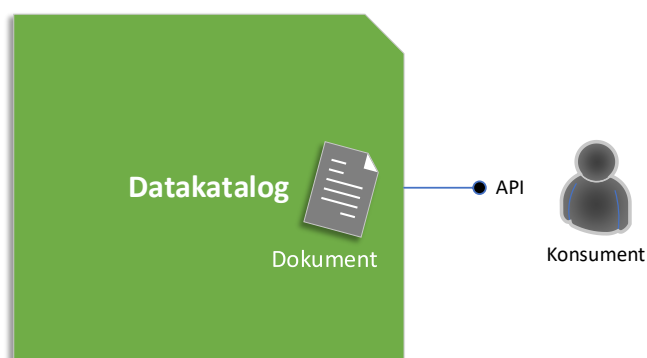


3.8 Specifikationer i datakatalogen

Hantering av specifikationer kan göras på olika sätt i en datakatalog och detta måste varje organisation själv bestämma vilket som passar dem bäst. Här redovisas tre olika sätt som kan användas. De tre sätten som redovisas är grundläggande och går att kombinera, men i detta dokument redovisar vi endast de grundläggande alternativen och vidareutveckling av alternativen kan respektive organisation kan ta vidare.

3.8.1 Specifikationer som dokument

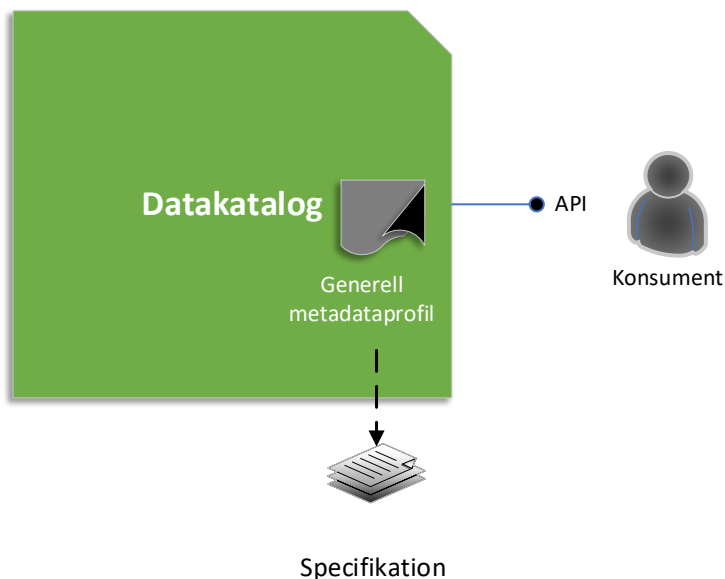
Ett alternativ är att själva dokumenten finns tillgängliga i datakatalogen konsument själv kan ladda ner dokumentet alternativt via en läsare läsa dokumentet direkt i datakatalogen.



Schematisk bild som beskriver att specifikationen finns i datakatalogen som ett dokument.

3.8.2 Specifikationer som generella metadataprofiler

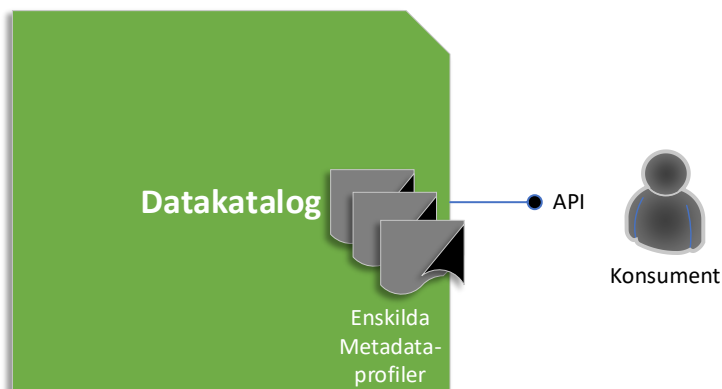
Ett alternativ är att det finns en generell metadataprofil som kan beskriva alla typer av specifikationer och via en länk ladda ner själva specifikationen. Den generella metadataprofilen stödjer alla typer av specifikationer och mer beskriver hur, var och när en viss specifikation ska användas, för att kunna bedöma vilken specifikation som konsumenten har behov av. I detta fall finns själva specifikationen inte i datakatalogen normalt sett.



Schematisk bild som beskriver hur en generell metadataprofil kan användas för att referera till specifikationer lagrade utanför en datacatalog.

3.8.3 Specifikationer som enskilda metadataprofiler

Alternativet med enskilda metadataprofiler används för att beskriva en specifikation i sin helhet och som i detta fall är likställt med dokumentet. De enskilda metadataprofilerna är då gjorda för exempelvis en Dataproduktspecifikation (DPS) eller för en API-specifikation. Detta gör att exempelvis en DPS beskrivs i sin helhet i metadataprofilen med metadata och på så sätt finns den även maskinläsbar och återfinns i själva datacatalogen.



Schematisk bild som visar hur enskilda metadataprofiler beskriver specifikationerna i en datacatalog.



4. Referenser

Nummer	Titel
1	ES2024-04 Styrning och ledningsfrågor av data (eSam)
2	ES2024 -09 Referensarkitektur: Datahantering och utbyte av data mellan myndigheter
3	ES2024-12 Gemensamma begrepp och kodlistor
4	ES2024 -11 Beskrivning och paketering av data som produkt
5	ES2023-01 AI – Utvecklingsprocessen och data (eSam)
6	ES2022-01 Vägledning Pseudonymisering av personuppgifter
7	FAIR Data Principles https://www.ccdc.cam.ac.uk/solutions/about-the-csd/fair-data-principles
8	Informationssäkerhet https://www.informationssakerhet.se/metodstodet
9	Vägledning för digital samverkan E-delegationen (eSam) https://www.digg.se/kunskap-och-stod/svenskt-ramverk-for-digital-samverkan
10	Begreppsmodell för metadatahantering version 1.5 https://www.naturvardsverket.se/data-och-statistik/tillgangliggorande-av-oppna-data/

eSam är ett medlemsdrivet program för samverkan mellan myndigheter för att underlätta och påskynda digitaliseringen inom det offentliga. eSam bildades 2015 som en frivillig fortsättning på E-delegationen. En viktig uppgift för eSam är att ta fram stöd och vägledningar som ger förutsättningar för att öka den digitala samverkan inom offentlig förvaltning.

Alla stöddokument finns på esamverka.se

I eSam ingår Arbetsförmedlingen, Arbetsmiljöverket, Bolagsverket, Boverket, Centrala Studiestödsnämnden, Domstolsverket, E-hälsomyndigheten, Ekonomistyrningsverket, Finansinspektionen, Folkhälsomyndigheten, Försäkringskassan, Havs- och vattenmyndigheten, Inspektionen för vård och omsorg, Jordbruksverket, Kemikalieinspektionen, Kriminalvården, Kronofogdemyndigheten, Kustbevakningen, Lantmäteriet, Länsstyrelserna, Migrationsverket, Naturvårdsverket, Patent- och Registreringsverket, Pensionsmyndigheten, Riksarkivet, Rättsmedicinalverket, Sida, Skatteverket, Skolverket, Statens institutionsstyrelse, Statens servicecenter, Statens tjänstepensionsverk, Statens veterinärmedicinska anstalt, Statistiska centralbyrån, Tillväxtverket, Trafikverket, Transportstyrelsen, Tullverket och Universitets- och högskolerådet (februari 2024)

