

Målarkitektur för öppna data

En generell beskrivning av vad en målarkitektur är och varför man ska ta fram en sådan, samt vilka tekniska förmågor som behöver finnas på etablerade för att en organisation ska kunna arbeta effektivt och kvalitativt med att tillgängliggöra öppna data

Creative Commons CC0

Innehåll

Inledning	3
Målgrupp.....	3
Metodik och upplägg.....	3
Avgränsningar	3
Om målarkitektur.....	4
Målarkitektur och dess relation till nuläges- och referensarkitektur.....	5
Målarkitektur i relation till stadsplan	6
Målarkitektur för öppna data.....	7
Antaganden som har gjorts	7
Omvärld och referensarkitektur.....	8
Datakonsumtionsmönster	8
Tekniska förmågor för öppna data	9
Metadatahantering	9
Datavisualisering	10
API-hantering.....	10
Integration och transformation	11
Databaslager	11
Datafilslager	11
Begreppsmodell.....	12
Egenskaper och övriga basförmågor.....	12
Säkerhet	12
Anpassningsbarhet	13
Underhållbarhet	13
Automatisering (för tillhandhållande av verksamhetsdata som öppna data)	13
Självetjäning (för tillhandhållande av verksamhetsdata som öppna data)	14
Övergripande användningsfall som målarkitekturen adresserar.....	14
Perspektiv: Dataproducent	14
Perspektiv: Datakonsument.....	14
Grundläggande principer för tillgängliggörande av data	15
Standarder.....	17
Datamodeller och format	18
Krav på IT-system och IT-tjänster som datakälla.....	18
Referenser och hänvisningar	19
Dokumenthistorik	20

Inledning

Syftet med målarkitekturdokumentet är dels att beskriva vad en målarkitektur är, dels att beskriva hur en framtida målarkitektur för öppna data bör vara uppbyggd inom en organisation. Målarkitekturen beskriver övergripande de IT-förmågor som är nödvändiga för att kunna arbeta effektivt och kvalitativt med att dela data öppet. Arkitekturen beskriver bland annat koncept och generella användningsfall som är relevanta för tillgängliggörande och publicering av öppna data.

I dokumentet har vi fokus på offentliga organisationer men givetvis är målarkitekturen även applicerbar på en privat organisation för samma ändamål. Det är också viktigt att poängtera att de förmågor som beskrivs i dokumentet även kan användas för andra IT-relaterade ändamål, dvs. inte enbart för att tillgängliggöra och publicera öppna data.

Dokumentet har ursprungligen tagits fram av Uppsala kommun (2021) och är till stora delar oförändrat. Vi har endast tagit bort de avsnitt som är unika för Uppsala kommun och ersatt vissa ord med mer generella begrepp samt kompletterat dokumentet med DIGG:s nationella principer för att tillgängliggöra information som öppna data. Dokumentet delas enligt licensformen Creative Common CC0 version 4.0.

Målgrupp

De primära målgrupperna för målarkitekturdokumentet är:

- CIO
- IT-arkitekter
- IT-infrastruktur och –systemtekniker
- IT-strateger
- Objektledare Verksamhet och IT
- Objektspecialister IT
- Objektägare Verksamhet och IT
- Verksamhetsutvecklare

Sekundära målgrupper är styrande och ledande forum inom organisationen, samt funktioner för juridik, informationsförvaltning, informations säkerhet samt registratur.

Metodik och upplägg

Arbetet med att ta fram en målarkitektur har bedrivits enligt följande metodik och upplägg:

- Genom omvärldsbevakning. Hur bygger andra offentliga och privata organisationer upp sin målarkitektur. Vilka referensramverk och standarder är relevanta att följa. Vilka tekniska lösningar finns på marknaden – nu och inom en snar framtid?
- Genom dialog med sakkunniga inom offentliga verksamheter.
- Genom att ta stöd av nationell expertis inom området.
- Genom praktiska exempel på delning av öppna datamängder.

Avgränsningar

Detta dokument beskriver inte i detalj hur målarkitekturen ska realiseras i praktiken. Den beskriver inte heller i detalj de ingående tekniska förmågorna.

Om målarkitektur

Det finns flera goda anledningar till att etablera en målarkitektur. En av dem är att målarkitekturarbete ökar chanserna att undvika onödig komplexitet i IT-landskapet, en annan är att främja flexibilitet och anpassningsbarhet över tid.

Poängen med att etablera en målarkitektur är att visa på en riktning baserat på långsiktiga mål, skapa förståelse inom sammanhanget och vara ett stöd i förflyttningen genom att vägleda och skapa medvetenhet i de val vi gör och varför vi gör dem. Med andra ord kan vi säga att vi ska sträva efter en avsiktlig arkitektur där vi säkerställer både interna och externa värden.

Dokumentet kan användas som grund för att utveckla respektive förmåga och detaljbeskriva i olika lösningsförslag. Arkitekturen kan fungera som underlag för uppdaterade anvisningar, arbetssätt, standarder samt lösningsarkitekturer för specifika verksamhetsförändringar som har behov av förmågor beskrivna i arkitekturen.

Med en målarkitektur avses följande:

- Beskrivningen av ett framtida tillstånd för den IT-arkitektur som utvecklas för en organisation.
- Målarkitekturen är (helt) anpassad efter organisationens och verksamhetens behov. Om den fanns i verkligheten så skulle den stödja organisationens behov på det mest optimala sättet. Den innehåller ett eller flera mållägen för IT-arkitekturen för ett avgränsat område eller sammanhang. En målarkitektur utgår ifrån "stadsplanen".
- Målarkitekturen beskrivs i hög utsträckning på konceptuell nivå. Det vi vet om det önskade IT-landskapet, även vissa detaljer, beskrivs i målarkitekturen. Detta medför att vissa delar i målarkitekturen är mer utvecklade och genomarbetade, medan andra är mindre utvecklade – och kan vara mer av en "platshållare".
- Målarkitekturen beskriver en mängd olika aspekter av IT-arkitekturer som speglar gjorda val och typer av möjliga lösningar. Perspektiven som beskrivs är utefter det sammanhang eller område som den berör. Målarkitekturen uttrycks med konceptuella modeller utifrån de olika perspektiven med bild och text. Den uttrycks också utifrån IT-arkitekturella principer som vägleder och/eller styr i vad som behöver följas eller uppnås.
- Målarkitekturens utformning baseras på kända styrsignaler i olika former inom organisationen. Dessa ger styrning och ledning i vad som beskrivs som viktigt, vilket beaktas under arbetet med framtagningen av målarkitekturen.

Målarkitekturen beskriver med andra ord ett önskat läge/tillstånd utifrån en helhet baserat på verksamhetens mål och vision samt ger vägledning i förflyttningen, dvs. den transformation som behöver genomföras. Det kan finnas flera framtida tillstånd presenterat i en färdplan för att visa utvecklingen av IT-arkitekturen till ett måltillstånd.

Målarkitekturen beskriver också viktiga förutsättningar, aspekter och områden som behöver adresseras och aktiviteter som ska genomföras för att framgångsrikt nå det önskade läget, utan att ge de exakta detaljerna.

Utöver att beskriva ett område och dess ingående delar så har arkitekturen till syfte att dels kunna användas som underlag i diskussioner med olika beslutsfattare, dels kunna användas vid t.ex. utvecklingsinsatser som framtagning av olika lösningar som ska stödja och vara en del av IT-arkitekturen inom kommunen.

Den ska kunna ge tillräckligt arkitekturstöd för planering, styrning och design och realisering av gemensamma förmågor och enskilda tillämpningar inom området som bidrar till en framåtriktad förflyttning mot målbilden.

Målarkitektur och dess relation till nuläges- och referensarkitektur

Kopplat till målarkitekturen kan referensarkitektur(er) användas och/eller utformas som kan tillämpas av verksamheterna med behov inom dessa områden, se *Bild 1*.

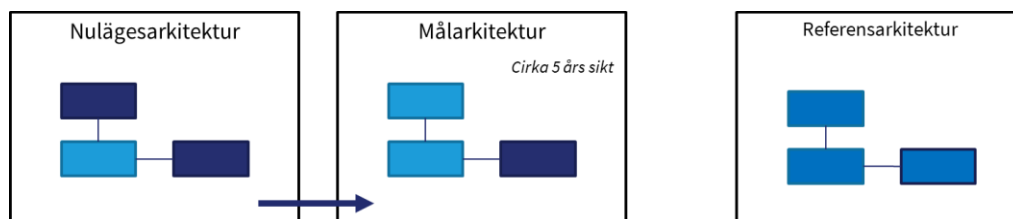


Bild 1 – illustration av hur nulägesarkitektur, målarkitektur och referensarkitektur förhåller sig till varandra

Målarkitekturen ska med andra ord bidra till att verksamheterna i en organisation, uppnår sina mål, kontinuerligt går i rätt riktning, förstår det sammanhang som berörs, samt får vägledning till ett enklare och snabbare beslutsfattande i och med att riktningen och vägen fram redan är utpekad.

En referensarkitektur används av en intressegrupp som en delad och överenskommen referensbeskrivning som kan användas för intressegruppens verksamhetsändamål. Referensarkitekturer används för att:

- hjälpa till att förstå formerna av troliga lösningar på problem inom en viss domän, och
- maximera möjligheten för gemensamma lösningsmönster för liknande problem inom en domän.

När beskrivningen är avsedd som en generell beskrivning som en guide för framtagande av en arkitektur, kallas den här typen av arkitektur ibland för en ”referensarkitektur”. I detta fall är inte själva referensarkitekturen avsedd att instansieras utan används som grund för att skapa en realiserbar arkitektur för någon enhet som är ett mer konkret exempel på den abstrakta referensarkitekturen.

Målarkitektur i relation till stadsplan

En målarkitektur tas i normalfallet fram för i ett specifikt syfte, för ett avgränsat verksamhets- eller teknikområde. Därför har en specifik målarkitektur alltid en specifik kontext att relatera till, se *Bild 2* för en illustration.

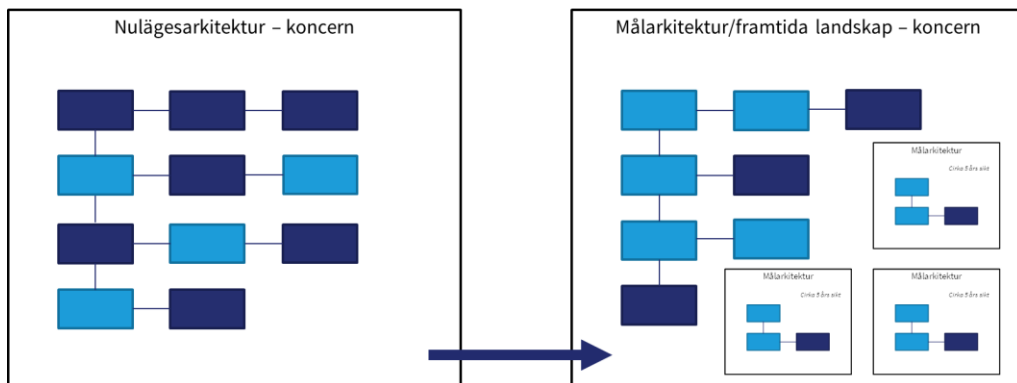


Bild 2 - illustration över hur målarkitektur förhåller sig till konceptet koncernarkitektur ("stadsplan")

Målarkitektur för öppna data

I detta avsnitt beskriver vi målarkitekturen för öppna data. I huvudsak fokuserar vi på de centrala och gemensamma förmågorna tillsammans med ett antal identifierade viktiga egenskaper för arkitekturen. En vidare beskrivning innefattar även de principer som gäller för området, samt användningsfall och standarder. För realisering av målarkitekturen ska även DIGG:s principer och vägledningar för att tillgängliggöra information ligga till grund.

För att kunna nå de nyttor och värden som finns i att arbeta med öppna data, enligt den framtagna målbilden, krävs IT-infrastrukturella förutsättningar i form av tekniska lösningar, arbetssätt och standarder. Målarkitekturen finns framtagen för att övergripande beskriva vad som behövs inom området för att kunna realisera en arkitektur som stödjer målbilden. Arkitekturen uttrycks i form av principer, standarder, konceptuella beskrivningar – som ska användas vid framtagande och utveckling av såväl gemensamma tekniska förmågor och funktioner som praktiska tillämpningar vid tillhandahållande/delning av öppna data.

Det centrala för en arkitektur inom öppna data innefattar de områden som berör de datamängder som är identifierade - med uthämtning av källdata, hantering av den, förberedning och slutligen tillhandahållandet.

Arkitekturen kommer att utvecklas och förädlas kontinuerligt i takt med att den används och kunskapen inom området ökas med hjälp av återkoppling baserat på den effekt den bidrar till samt vid omvärldsbevakning inom området avseende exempelvis möjliggörande teknik, metoder, standarder och krav.

Sannolikt finns redan ett flertal av de tekniska förmågor som vi beskriver i denna målarkitektur redan på plats och används för andra ändamål inom organisationen. Med andra ord är inte förmågorna unikt knutna till att enbart tillgängliggöra öppna data.

Antaganden som har gjorts

Målarkitekturen har arbetats fram med utgångspunkt från följande antaganden:

- Krav på tekniska funktioner inom öppna data förväntas över tid att vara likvärdiga från organisationens olika verksamheter – och kan därmed tillgodoseas i en gemensam teknisk lösning.
- Verksamheterna som ska publicera öppna data (dataproducenterna) har likvärdiga tekniska behov, även om sättet som de öppna datamängderna delas på skiljer sig åt.
- Där organisationens verksamheter har likartade gemensamma behov är utgångspunkten att nyttja gemensamma tekniska lösningar, förutsatt att respektive verksamhetsbehov kan omhändertas och bemötas tillräckligt.
- Tekniska lösningar för organisationens behov utgår ifrån att vara centraliserade – i formen av tjänster och funktioner som kan nyttjas av de olika verksamheterna.
- Vid behov av unika uppsättningar av tjänster och funktioner, för att möta verksamhetens specifika behov eller där divergens är önskvärd, sker en värdering som tar in både gemensamma och individuella verksamhetens behov och önskemål.

- De öppna datamängderna och de olika interna och externa målgrupperna (datakonsumenterna) kommer att ställa olika krav på hur data delas både avseende teknisk lösning och format, vilket målarkitekturen ska stödja.
- Datakonsumenter har behov av olika färskhet på data, ex statiska data och realtidsdata. Därför behöver en framtida målarkitektur ha stöd för publicering av såväl statiska data som strömmande realtidsdata – och allt däremellan.
- Behoven relaterat till att dela data kommer att förändras över tid och visa sig tydligare när delande av data sker i större utsträckning än idag, vilket kommer att leda till mer insikter och krav på den tekniska plattformen och därmed tekniska arkitekturen.

Omvärld och referensarkitektur

Vid framtagningen av detta dokument har omvärldsbevakning genomförts med syftet att hitta olika exempel på målarkitekturer för öppna data. Det har visat sig vara en utmaning. Vi har tagit del av Stockholm stads målarkitekturdokument för sakernas internet (IoT) samt Helsingborgs tekniska arkitektur för deras så kallade datavalv. Utöver det har vi även tagit del av tekniska skisser på liknande arkitekturer som Stockholm och Helsingborg, som omfattar mer än bara hantering av öppna data. Avslutningsvis har vi skickat detta målarkitekturdokument på remiss till medverkande regioner, kommuner och leverantörer i det så kallade NSÖD-projektet (Nationell Skalning Öppna Data) för att få in kloka synpunkter, förslag och eventuella justeringar av dokumentet.

Datakonsumtionsmönster

Datakonsumenterna, dvs. de som vill nyttja öppna data, kommer att ha olika behov och förväntningar på den öppna data som verksamheterna publicerar över tid.

Det kan exempelvis handla om datamängdernas aktualitet, riktighet och tillgänglighet (servicenivå). För historisk forskning är kanske datamängdens aktualitet inte så viktig medan dess riktighet är betydligt viktigare. Omvänt kanske inte datamängdens hundraprocentiga riktighet är lika viktig om vi placerat ut 100 relativt billiga sensorer som mäter partikelutsläpp runt om i staden, utan det viktigaste i detta fall kanske är att vi kan samla in en stor mängd realtidsdata för att se skiftningar i partikelutsläpp under en dag, en vecka, en månad eller ett år. Vi kan leva med att sensorerna har viss feltolerans, antalet mätpunkter är viktigare.

Datakonsumenterna kommer förvänta sig att det finns olika sätt att ta del av den öppna data som publiceras. Vissa kommer vilja ladda ner öppna data som filer medan andra datakonsumenter vill anropa öppna data maskinläsbart genom så kallade maskinläsbara gränssnitt (API, Application Programming Interface), direkt från de applikationer som de utvecklar eller använder för sina olika syften. En tredje kategori kanske endast vill ta del av öppna data genom en datavisualisering (karta, diagram osv) som de interna verksamheterna eller en annan extern datakonsument skapat.

På motsvarande sätt kommer datakonsumenterna förvänta sig att öppna data publiceras i format och standarder som möjliggör effektiv och enkel återanvändning. En förutsättning för detta är att det finns en bra dokumentation som beskriver vad de öppna datamängderna innehåller för data, hur datafälten ska tolkas, datafältens inbördes relation med mera.

I den nya öppna datalagen lyfts det särskilt fram att myndigheter (dvs. stat, region och kommun) i så stor utsträckning som möjligt ska tillmötesgå datakonsumenternas önskemål om format. Det framgår också att en styrning mot maskinläsbara och sökbara format förordas och att särskild värdefulla och dynamiska data ska tillhandahållas via lämpliga gränssnitt och, om det är möjligt, för bulknedladdning

Målarkitekturen ska stödja alla dessa behov. Statiska data, data som uppdateras med visst intervall, realtidsdata, data som är åtkomlig via både maskinläsbara gränssnitt och olika filformat och standards.

Tekniska förmågor för öppna data

Relationen mellan användare och de tekniska förmågor som en arkitektur för öppna data behöver omfatta, kan schematiskt beskrivas enligt grafiken nedan. Förmågorna i sin helhet finns till för att möjliggöra en effektiv hantering inom den interna organisationen och en god upplevelse för datakonsumenterna, där förmågorna ska kunna interagera på ett effektivt och funktionellt sätt med sin omgivning.

Syftet med uppdelningen är att dels skapa oberoende mellan de ingående delarna, dels förståelse för respektive skikt och därmed ansvar vid delning respektive konsumtion av data.

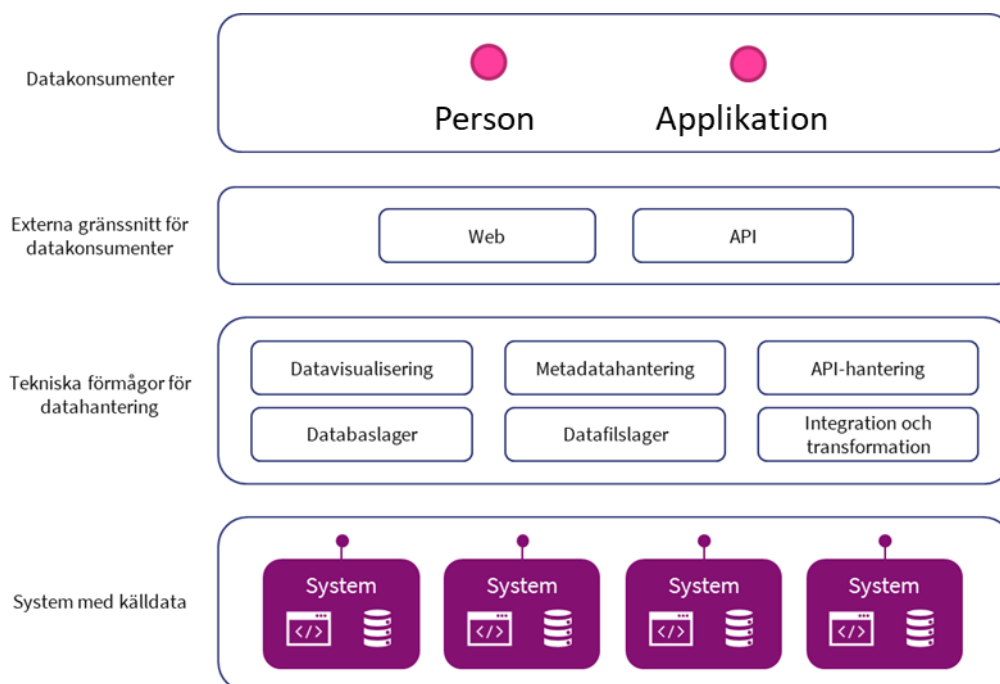


Bild 3 – illustration över tekniska förmågor som behöver realiserats för öppna data

På kommande sidor beskrivs de tekniska förmågorna för datahantering mer i detalj, dvs. för vilket syfte de är en komponent i målarkitekturen.

Metadathantering

Metadata betyder rent språkligt data om data eller information om data. Normalt används metadata för att beskriva innehållet och/eller strukturen för en viss datasamling ur olika perspektiv. En av de vanligaste tillämpningarna är inom dokumentation där metadata används i dokumentshuvuden för att ge svar på sådant som dokumentets författare, namn, titel, identitet, utgåva och datum.

En metadatakatalog används i detta sammanhang för att beskriva den öppna data som publiceras i katalogen. Syftet är dels att data ska vara sökbar och länkbar på Internet, dels att användaren (datakonsumenten) ska förstå vad datamängden innehåller och hur den ska tolkas. Metadatakatalogen används för att publicera öppna datamängder i form av hänvisningar till filer eller API:er med tillhörande dokumentation.

För att publicerade datamängder ska kunna visas automatiskt på Sveriges och EU:s dataportaler behöver metadatakatalogen stödja EU-standarden DCAT-AP (Application Profile for data portals in Europe) fullt ut.

Katalogen ska även ha stöd för att hantera vid var tids gällande standarder avseende filformat (ex .csv och XML) och API-anrop (ex. JSON). Det kan även finnas behov av att stödja andra typer av standards på sikt, vilket leverantören ska vara garant för genom klausuler i licensavtalet.

Katalogen ska även ha stöd för att läsa in och publicera filer utan att annan lagringsyta ska krävas. Katalogen bör dessutom ha stöd för att dela interna öppna data, dvs. data som inte är åtkomliga för externa datakonsumenter.

Datavisualisering

För att öka förståelsen om den öppna data som publiceras, kan det vara bra att visualisera data i en datavisualiseringsplattform. Plattformen bör ha bra funktioner för att bryta ner och beskriva den visualiserade datamängden på ett lättförståeligt sätt för datakonsumenten i form av grafer, kartor och så vidare. Plattformen bör även ha funktioner för att datakonsumenten ska kunna återanvända visualiseringen på ett enkelt sätt i andra system och plattformar, som exempelvis på en webb (HTML, Javascript osv.).

API-hantering

Ett API är ett applikationsprogrammeringsgränssnitt som kan användas av ett program för att kommunicera med ett annat program eller system, exempelvis för att utbyta information och/eller anropa olika former av funktioner i ett annat system. Det finns olika typer av API, det vi fokuserar på i detta sammanhang är API:er som används av ett program för att läsa information i form av öppna data från ett underliggande IT-system som används av organisationens verksamhet(er) (dataproducenten).

En API-plattform kan i detta sammanhang användas för att skapa ett så kallat tjänstelager och en gateway mellan datakonsumentens program och det underliggande IT-systemet där den öppna datamängden finns lagrad. På så sätt skyddas det underliggande IT-systemet av API-plattformen. Dessutom skapas det förutsättningar för att bygga ett anpassat API (för den öppna datamängden) som anropar ett standard-API i det underliggande IT-systemet, för att hämta ut den information som är intressant i sammanhanget.

Det är inte nödvändigt att anskaffa en hel API-plattform, det går även att utveckla enskilda tjänster som körs på en server för att svara på anropen.

JSON (JavaScript Object Notation) är ett format som är vanligt förekommande för att anropa öppna data via ett API.

Integration och transformation

En integrationsmotor används för att läsa ut, anpassa och skriva information från och till olika IT-system (läs databaser) och filer. En integrationsmotor har även inbyggda funktioner för att skicka och ta emot filer via olika typer av protokoll som exempelvis webservices (REST API/SOAP), e-post (SMTP/POP3), SFTP (Secure File Transfer Protocol), EDI (Electronic Data Interchange) eller krypterat via Internet, för att nämna några av de vanligast förekommande protokollen.

I en integrationsplattform finns det även möjlighet att bygga in logik för att exempelvis tolka, filtrera, konvertera och komplettera in-data på olika sätt för att passa det mottagande systemet – i detta sammanhang den öppna datamängden som ska publiceras i metadatakatalogen. Exempelvis kan integrationsplattformen hämta ut grundläggande information om kommunens leverantörsfakturer och spara dessa i en fil (exempelvis i Excel CSV-format) som sedan publiceras som öppna data i metadatakatalogen. Eftersom vi inte vill publicera personuppgifter så filtreras alla enskilda firmor bort från urvalet genom logik som utvecklats i integrationsmotorn. Dessutom filtreras alla fakturer med sekretessmärkning bort från urvalet av samma anledning.

Databaslager

Databaslagret används för att lagra och aggregera data från verksamheternas olika IT-system och publicera dem som öppna data i metadatakatalogen. Ett databaslager av denna typ placeras på DMZ (DeMilitarized Zone).

Databaslagret används exempelvis när det inte är lämpligt att exponera de interna IT-systemen (databasen) mot Internet samt när man vill aggregera data från flera olika datakällor till en och samma datamängd. Den data som lagras kan därefter göras tillgänglig för potentiella datakonsumenter via ett API (se tidigare avsnitt). Vanligtvis sker anropen via ett API som anropar en vy i databasen för att hämta och presentera den data som efterfrågas av datakonsumenten.

Exempelvis kan det vara relevant att läsa ut grundläggande information (ärendemening, nämnd, datum osv.) om kommunens nämndärenden från kommunens diarium till detta databaslager och publicera som öppna data i metadatakatalogen. Datakonsumenterna kan därefter anropa och söka bland alla nämndsärenden i databaslagret utan att kommunens diarium påverkas.

Datafilslager

Datafilslagret används för att lagra olika former av filer som ska publiceras som öppna data. Även om metadatakatalogen har funktioner för att lagra filer som ska publiceras som öppna data, kan det vara relevant att använda ett separat datafillager av olika anledningar.

Exempelvis kan det underlätta administrationen att regelbundet och automatiserat läsa ut filer till ett datafilslager istället för att läsa in dem manuellt i metadatakatalogen. Ett annat exempel är om det är en stor mängd filer eller stora filer som ska publiceras som öppna data, som ställer krav på stor lagringskapacitet. I dessa fall kan det vara mer ekonomiskt fördelaktigt att använda ett internt datafilslager, istället för att lagra den öppna datamängden i metadatakatalogen, speciellt om metadatakatalogen är en molnbaserad tjänst. Informationssäkerheten är en tredje aspekt.

Tillgänglighet

För datakonsumenten är det viktigt att veta vilken tillgänglighet den öppna datamängden har. Arkitekturens tillgänglighet behöver vara förenlig med karaktären på de datamängder som delas och datakonsumenternas förväntningar på tillgänglighet. Exempelvis så kan en hög nyttjandegrad av en datamängd med hög periodicitet (realtid eller nära realtid) som delas med hjälp av dataströmmar ställa stora krav på tillgänglighet.

En stabil webbadress varifrån informationen kan skördas, inklusive robust infrastruktur för namnuppslag (DNS), samt unika och beständiga identifierare, även kallad persistent identifierare (PID) ska användas i så stor omfattning som möjligt.

Spårbarhet

Det är värdefullt och i viss mån nödvändigt att genom arkitekturen kunna spåra var anrop härstammar från och vilken data som anropas, i syfte att kunna stödja både användningen av data och i felsöknings syfte. En gemensam loggfunktion för teknisk spårbarhet är nödvändigt för att stödja detta.

Anpassningsbarhet

Arkitekturen behöver kunna leverera sina tjänster och sitt värde effektivt med god upplevelse över tid, i takt med att volymen på de öppna datamängderna och datakonsumtionen varierar, inklusive olika sätt och mönster i hur datamängderna delas. Både vid låga och höga volymer av öppna datamängder samt vid låg och hög användning, konstant och periodiskt, så ska en god upplevelse säkras för datakonsumenterna. Detta ska helst kunna säkerställas utan återkommande justering av målarkitekturens ingående komponenter. En inbyggd elasticitet behövs således finnas inom arkitekturen.

Underhållbarhet

Arkitekturens ingående delar ska i praktiken vara förhållandevis löst kopplade för att kunna förändras och bytas ut över tid, med minimal inverkan på övriga komponenter i arkitekturen och arkitekturens övergripande funktionalitet och kvalitet.

Automatisering (för tillhandhållande av verksamhetsdata som öppna data)

Arkitekturen ska stödja en hög grad av automatisering, dels för att möjliggöra en god och ändamålsenlig service och datakvalitet, dels för att minska behovet av löpande administration och risken för mänskliga handhavandefel.



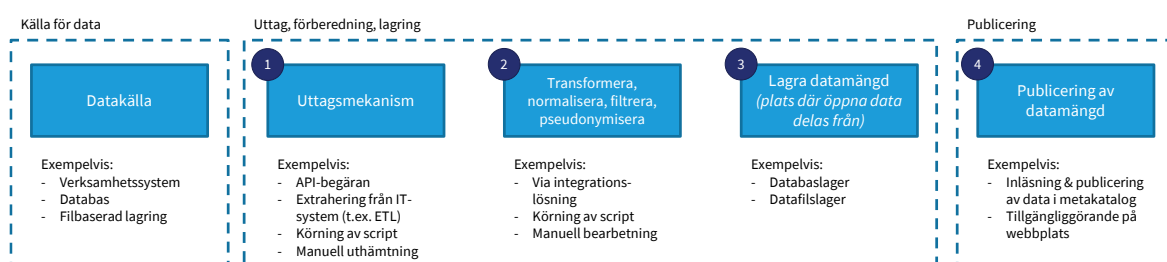
Självbetjäning (för tillhandhållande av verksamhetsdata som öppna data)

Arkitekturen ska stödja en nivå av självbetjäning som stimulerar delning av öppna data, där de interna arbetsuppgifterna ska vara enkla, självförklarande och ej kräva djup teknisk kompetens att utföra hos användaren som delar data.

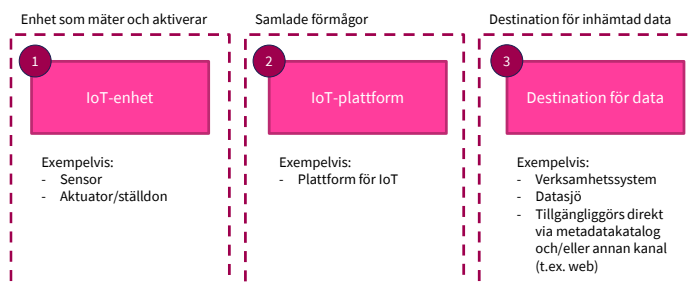


Övergripande användningsfall som målarkitekturen adresserar

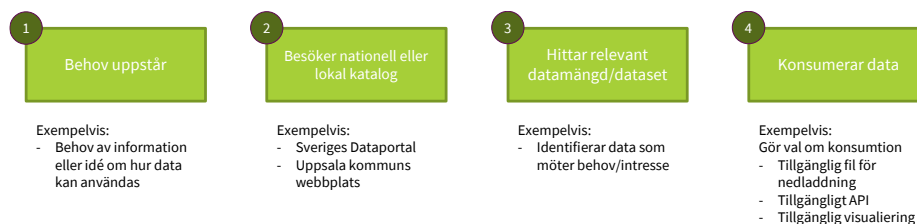
Perspektiv: Dataproducent



I fallet med realtidsdata baserat på IoT-tillämpning. Notera att målarkitekturen för öppna data inte omfattar förmågor för att hantera IoT-enheter.



Perspektiv: Datakonsument



Grundläggande principer för tillgängliggörande av data

Arbetet med att tillgängliggöra och publicera öppna data bör baseras på DIGG:s nationella principer för att tillgängliggöra information. Principerna syftar till att bidra med konkretisering och exempel på vilka krav som målarkitekturen ska leva upp till. Principerna ska följas i största möjliga utsträckning för att säkerställa att utvecklingen arkitekturen för att kunna genomföra den nödvändiga förflyttningen inom området. Dock kan avvikelser från dessa principer göras utifrån specifika behov av tillämpningar inom arkitekturen, vilket behöver göras medvetet och kontrollerat.

I kommande avsnitt har vi kopierat in de nationella principerna för att tillgängliggöra information, samt kompletterat med kommentarer i de fall det är relevant.

Princip 1: Öppenhet som standard

Verksamheten ska göra information öppet och digitalt tillgänglig för användning av externa intressenter, på det sätt som beskrivs i efterföljande principer. Endast den information som är skyddsvärd ska särbehandlas.

Detta arbete ska ses som en del i att bidra till att data och information blir en bred värdeskapande resurs för samhället som bidrar till innovation, transparens, kunskap och effektivitet. Att skapa en öppen och tillgänglig dataresurs från offentliga organisationer kräver därför ett kontinuerligt arbete där tillgängliggörande av information är en naturlig del av verksamheten och dess kultur.

I detta sammanhang är det viktigt att de it-system och digitala tjänster som löpande upphandlas lever upp till denna princip. I praktiken innebär det att säkerställa att organisationen äger den data som skapas och har tillgång till den över tid, exempelvis när ett it-system eller en digital tjänst ska fasas ut/ersättas. Lika viktigt är det att ställa krav på att det finns öppna API:er för informationsutbyte med andra interna och externa it-system och digitala tjänster.

Princip 2: Bedriv ett riskbaserat och systematiskt informationssäkerhetsarbete

Informationen som tillgängliggörs ska vara informationsklassad utifrån aspekterna konfidentialitet, riktighet och tillgänglighet för att identifiera konsekvenser av ett otillräckligt skydd vid och under tillgängliggörandet. I sin riskbedömning ska verksamheten beakta att den tillgängliggjorda informationen kan ackumuleras och aggregeras med annan information vilket kan ge upphov till skador för exempelvis individers integritet, eller organisationer och samhällets säkerhet.

Återkommande informationsklassning och riskbedömning ska ligga till grund för beslut om publicering och lämpliga säkerhetsåtgärder under publicerandet. Verksamheten bör säkerställa att säkerhetsåtgärder kring tillgängliggörandet anpassas efter eventuella förändringar i behov, användningsområden och teknisk utveckling.

Exempelvis är det viktigt att bedöma om källsystemet, dvs. det it-system eller den digitala tjänst där datamängden lagras, ska exponeras direkt mot internet eller om denna data ska läsas ut och exponeras på annat sätt enligt de exempel som beskrivits tidigare i avsnittet Tekniska förmågor för öppna data.

Princip 3: Tillgängliggör aktuell och uppdaterad information med användaren i centrum

Verksamheten ska sträva efter att information som tillgängliggörs ska vara aktuell och innehålla så mycket som möjligt till exempel när det gäller detaljeringsgrad och omfattning över tid.

Verksamheten ska samtidigt sträva efter att anpassa innehåll, format, servicenivåer och dokumentation utifrån användarnas behov för att öka möjligheterna till att informationen används. Tillgängliggörandet ska ske utan onödig fördröjning och bör inte undanhållas för att informationen inte är fullständig, brister i kvalitet eller att efterfrågan ännu inte identifierats.

Den nya öppna datalagen (2022: 818) som syftar till att främja den offentliga sektorns tillgängliggörande av data för vidareutnyttjande, särskilt i form av öppna data, ställer krav på att offentliga verksamheter. Vid en begäran ska en offentlig verksamhet kunna lämna ut den begärda datan inom fyra veckor. Det finns även möjlighet att begära dispens på ytterligare fyra veckor om särskilda behov föreligger.

Lagen säger att denna data i första hand ska lämnas ut i befintligt format, för att undvika kostnader med att konvertera denna data till olika format. Samtidigt uppmuntrar lagen offentliga verksamheter att arbeta proaktivt med att publicera öppna datamängder som ofta efterfrågas. Likaså att publicera dessa data med den kvalitet och i de format som efterfrågas, samt maskinläsbart i de fall det är möjligt, vilket även är relevant för princip 4.

Princip 4: Gör information lätt att hantera

Verksamheten ska sträva efter att använda format, API:er och specifikationer som är öppna, standardiserade och maskinläsbara. Det ökar möjligheterna för att samverka digitalt, kombinera och bearbeta information och lägger inte tekniska begränsningar på hur informationen kan användas.

Verksamheten ska sträva efter att använda öppna och kontrollerade begrepps- och kodlistor samt beständiga identifierare eftersom information får ett mervärde genom att hänvisa till samma plats, ämne eller händelse.

Princip 5: Använd villkor som främjar bred användning

Information ska publiceras under villkor som inte i onödan begränsar möjligheterna till vidareutnyttjande. Verksamheter ska som utgångspunkt inte ställa krav på att användaren behöver ansöka om tillstånd eller registrera sig. Om villkor sätts upp ska dessa vara objektiva, proportionella, icke-diskriminerande och motiverade av ett allmänintresse. Rättighetsmärkningarna och licenser som används ska vara öppna, maskinläsbara och vedertagna.

Huvudregeln är att information ska erbjudas avgiftsfritt om det inte framgår annat i lag, förordning, eller särskilt beslut av regeringen. Om avgifter tas ut ska beräkningsgrunden publiceras öppet och elektroniskt.

Princip 6: Dokumentera och beskriv information

Information som tillgängliggörs ska dokumenteras och beskrivas så att den är lätt att upptäcka, förstå och användas av både människor och maskiner. Metadata om den information verksamheten tillhandahåller digitalt för vidareutnyttjande ska vara sökbar på den nationella dataportalen.

Annan dokumentation kan innehålla kodexempel, versionshantering, beskrivning av hur informationen skapats, vad den används till idag och exempel på användning. Kvaliteten på uppgifterna ska också dokumenteras och kända felaktigheter och utmaningar ska uttryckligen beskrivas.

Det är också relevant att beskriva sina metadata och sin dokumentation på engelska eftersom det kan finnas datakonsumenter som inte kan läsa och förstå svenska, exempelvis inom näringslivet och akademien.

Princip 7: Uppmuntra till användning och dialog

Verksamheten ska sträva efter att interagera med användare och aktivt uppmuntra till användning av den tillgängliggjorda informationen. När verksamheten själv använder informationen ska de sträva efter att använda den information som publicerats.

Verksamheten bör också erbjuda användarna ett aktivt stöd som möjliggör återkoppling, dialog och frågor, och som kännetecknas av snabbhet och upplevd kvalitet. Denna interaktion skapar bättre möjligheter för verksamheter att fånga upp olika behov och att förbättra kvaliteten på den tillgängliggjorda informationen, samt att det skapar ökade möjligheter för vidareutnyttjare att skapa nya värden och innovationer.

Standarder

I detta avsnitt beskrivs ett antal standarder som är vanligt förekommande inom öppna data och är därmed bra eller nödvändiga att känna till.

Namn	Beskrivning
DCAT-AP	DCAT-AP-SE är en metadata-specifikation som syftar till att beskriva datamängder. Mer specifikt är DCAT-AP-SE en svensk anpassning (en så kallad profil) av den europeiska metadata-specifikationen DCAT-AP. DCAT-AP är i sin tur en anpassning (profil) av den internationella rekommendationen DCAT som utges av W3C. Läs mer: DCAT-AP-SE version 2.0.0 (dataportal.se)
JSON	JSON, är ett kompakt, textbaserat format som används för att utbyta data. Formatet är utformat för att enkelt fungera med Javascript, men i dag kan nästan alla programspråk skapa och läsa JSON-kod. Det är enkelt för både människor och applikationer att läsa, skriva och exekvera.
CSV	CSV står för Comma Separated Values och avser textfiler som innehåller data, där datafälten är separerade med exempelvis ett kommatecken. Även andra tecken kan användas som separator, exempelvis ett semikolon.
WMS	Web Map Service är ett protokoll för att publicera georefererad kartinformation i rasterformat över internet. Protokollet är en standard framtagen av Open Geospatial Consortium.
XML	XML är en förkortning som står för Extensible Markup Language. Det är ett standardiserat språk för att bygga upp olika typer av dokument samt att strukturera informationen med en tydlig märkning. Med märkning menas att man i texten lägger in förklarande markörer, precis som man gör när man skapar webbsidor. Med XML kan man med andra ord skapa ett standardiserat beskrivningsspråk för att beskriva en definierad typ av information, exempelvis en öppen datamängd.

Datamodeller och format

Den öppna data som verksamheterna löpande kommer att publicera, kommer att vara av olika karaktär och ha olika interna och externa målgrupper (datakonsumenter), som i sin tur kommer att ha olika skäl och syften till varför de vill använda den öppna data som tillgängliggjorts.

För varje öppen datamängd som ska publiceras kommer verksamheterna att behöva ta ställning till hur datamodellen ska se ut och i vilket/vilka format som datamängden ska publiceras. Verksamheterna bör alltid sträva efter att använda datamodeller och format som andra kommuner, regioner, statliga myndigheter eller branscher använder, eftersom det ökar värdet och underlättar återanvändningen av data för datakonsumenten. Om exempelvis många kommuner publicerar exakt samma typ av datamängd enligt samma datamodell och format så är det sannolikt ännu mer intressant för en datakonsument att använda den data som publicerats.

I de fall en sådan standardiserad dataspecifikation saknas, kan det vara relevant att ta fram en egen specifikation enligt tidigare nämnda principer för tillgängliggörande av data. På den nationella dataportalen finns det ett flertal dataspecifikationer publicerade som kan användas som inspiration i detta sammanhang. På dataportalen finns det även en Community som kan användas som stöd i detta arbete.

Krav på IT-system och IT-tjänster som datakälla

IT-system och digitala tjänster som verksamheterna anskaffar eller utvecklar själva ska som regel leva upp till DIGG:s nationella principer för att tillgängliggöra information samt rekommendationer för att upphandla data, vilket ska garantera ägandeskap och tillgång till verksamheternas data över tid, samt möjlighet att utbyta data med andra aktörer enligt öppna standards och format.

Referenser och hänvisningar

I följande tabell finns referenser och hänvisningar till externa källor och externa resurser som nämns i dokumentet.

Nationella principer för att tillgängliggöra information (DIGG)	https://www.digg.se/utveckling-av-digital-forvaltning/oppna-och-delade-data/offentliga-aktorer/nationella-principer
Vägledning för att tillgängliggöra information (DIGG)	https://www.digg.se/utveckling-av-digital-forvaltning/oppna-och-delade-data/offentliga-aktorer/vagledning-tillgangliggora
Rekommendationer för upphandling av data (DIGG)	https://www.digg.se/kunskap-och-stod/oppna-och-delade-data/offentliga-aktorer/rekommendationer-for-upphandling-av-data
Vägledning om begäran för vidareutnyttjande av data (DIGG)	https://www.digg.se/kunskap-och-stod/oppna-och-delade-data/offentliga-aktorer/information-om-oppna-datalagen/om-begaran-om-tillgangliggorande-av-data-for-vidareutnyttjande
Sveriges dataportal	https://www.dataportal.se
EU:s dataportal (European Data portal / The official portal for European data)	https://data.europa.eu/en
DCAT Application Profile for data portals in Europe	https://joinup.ec.europa.eu/collection/semantic-interoperability-community-semic/solution/dcat-application-profile-data-portals-europe
NSÖD (Nationell Skalning Öppna Data)	https://community.dataportal.se/topic/169/projektet-ns%C3%B6d-nationell-skalning-%C3%B6ppna-data
Målarkitektur IoT och dataplattform (Stockholms stad)	https://smartstad.stockholm/wp-content/uploads/sites/10/2021/05/M%C3%A5larkitektur-f%C3%B6r-IoT-och-dataplattform.pdf

Dokumenthistorik

Version	Utförda förändringar	Skapat datum	Skapad av	Godkänt datum	Godkänd av
0	Grovt utkast på innehållsstruktur	2021-09-01	PB		
1	Flyttat över innehåll från PPT till detta dokument utan några direkta tillägg eller justeringar	2021-10-19	PB		
2	Justeringar av struktur och innehåll	2021-10-19	PB		
3	Kompletterat med information där det saknades i vissa avsnitt och rensat en del	2021-11-05	FE		
4	Nya avsnitt, utvecklade avsnitt, kompletteringar samt justeringar	2021-11-17	MP		
5	Justeringar och kompletteringar av samtliga avsnitt	2021-12-01	MP		
6	Mindre justeringar och strukturändringar	2021-12-02	MP		
7	Mindre justeringar. Kommentarer på principer.	2021-12-06	PB, FE		
8	Samtliga principer uppdaterade.	2021-12-13	MP		
9	Slutliga justeringar innan utskick för remiss.	2021-12-16	MP		
10	Alla anknytningar till Uppsala kommun har tagits bort. Dokumentet är mer allmänt skrivet. Avsnittet om principer följer nu de nationella principerna	2023-02-20	FE		